

# 十分钟智商运动

李永乐◎著

十分钟的时间，做些有趣的大脑运动。




**博学和风趣亦能兼得，学识和学历可以无关**

**用知识武装头脑，在人群中闪闪发光**

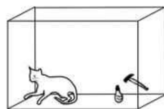
清华北大 + 人大附中 } 乐学呆萌天才

西瓜视频 + 新浪微博 } 百万明星博主

**“国民老师”李永乐 首部趣味科普书**








百花洲文艺出版社
BAIHUAZHOU LITERATURE AND ART PRESS

书籍朋友圈分享微信Booker527

VISIT...

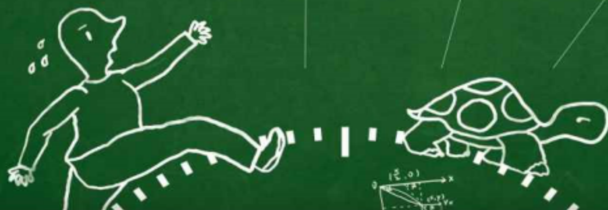
LANZAROTE Caliente.COM




# 十分钟智商运动

李永乐◎著

十分钟的时间，做些有趣的大脑运动。

**博学和风趣亦能兼得，学识和学历可以无关**

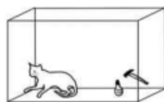
**用知识武装头脑，在人群中闪闪发光**

清华北大 + 人大附中 } 乐学呆萌天才

西瓜视频 + 新浪微博 } 百万明星博主

**“国民老师”李永乐 首部趣味科普书**



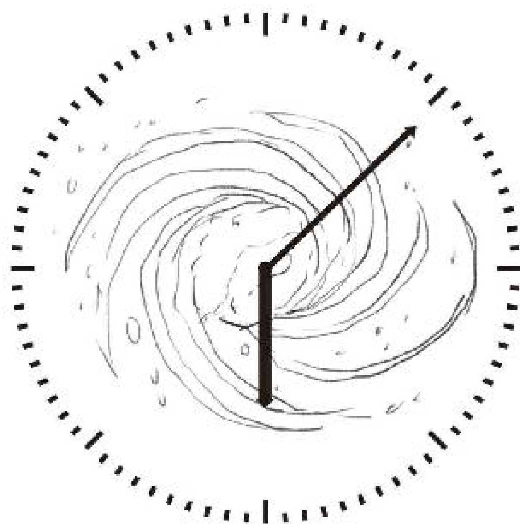




百花洲文艺出版社
BAIHUAZHOU LITERATURE AND ART PRESS

书籍朋友圈分享微信Booker527

# 十分钟智商运动

李永乐◎著



百花洲文艺出版社

图书在版编目（**CIP**）数据

十分钟智商运动 / 李永乐著．—南昌：百花洲文艺出版社，2018.12

ISBN 978-7-5500-3096-1

Ⅰ．①十…　Ⅱ．①李…　Ⅲ．①科学知识—普及读物　Ⅳ．①Z228

中国版本图书馆CIP数据核字（2018）第247667号

十分钟智商运动

SHI FENZHONG ZHISHANG YUNDONG

李永乐　著

出 版 人　姚雪雪

出 品 人　李国靖

特约监制　陈美珍

责任编辑　游灵通　程　玥

特约策划　刘孟琳

特约编辑　刘孟琳

封面设计　A BOOK STUDIO 萝卜Design 1092801781

版式设计　赵梦菲　阅众时刻

内插绘图　谢　嘉　刘　恒

出版发行　百花洲文艺出版社

社　　址　南昌市红谷滩世贸路898号博能中心I期A座20楼

邮　　编　330038

经　　销　全国新华书店

印　　刷　嘉业印刷（天津）有限公司

开　　本　710mm×1000mm　　1/16

印　　张　15.25

字　　数　250千字

版　　次　2018年12月第1版第1次印刷

书　　号　ISBN 978-7-5500-3096-1

书籍朋友圈分享微信**Booker527**

版权所有，侵权必究

发行电话　0791-86895108

网　址　http://www.bhzwy.com

图书若有印装错误，影响阅读，可向承印厂联系调换。

# 目录

# contents


第一章　有趣的数学

1.1　最早的学霸是谁？

1.2　阿基里斯能追上乌龟吗？

1.3　我说的是假话

1.4　3.1415926...

1.5　比萨饼中的数学

1.6　一笔能写出“田”字吗？

1.7　1 + 1

1.8　最厉害的民科是谁？

1.9　怎样传小纸条？

1.10　平行线存在吗？

1.11　四维空间是什么样的？

1.12　数学家能在赌场中赢钱吗？

1.13　买彩票能中大奖吗？

1.14　天气预报为啥总不准？

1.15　散户炒股为啥总赔钱？

1.16　老板为啥对底层员工特别好？

1.17　为啥总有人开车加塞？

第二章　奇妙的物理

2.1　能量都是从哪儿来的？

2.2　光速是如何测量的？

2.3　阿基米德能撬起地球吗？

2.4　天体之间的距离到底有多远？

2.5　指南针为啥能指南？

2.6　家用电是怎么产生的？

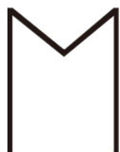
2.7　特斯拉和爱迪生谁更厉害？

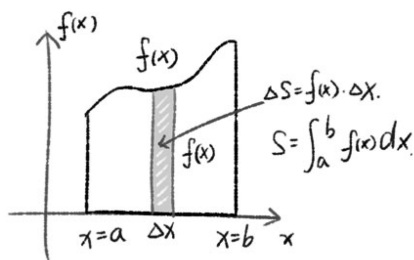
2.8　SOS是什么意思？

2.9　FM和AM啥意思？
2.10　世界上第一张X光片是谁拍的？
2.11　量子是个什么玩意儿？
2.12　波还是粒子？这是个问题
2.13　薛定谔的猫是死了还是活着？
2.14　黑洞是黑色的吗？
2.15　如何制作原子弹
2.16　芯片为啥这么难做？
第三章　身边的科学
3.1　天空为什么是蓝色的？
3.2　星星为什么是黑白的？
3.3　正常夫妻为啥会生出色盲孩子？
3.4　双层彩虹是怎么回事？
3.5　炎热的夏天为啥总感觉马路上有水？
3.6　雨中走路淋雨多，还是跑步淋雨多？
3.7　电磁炉是怎么加热食物的？
3.8　微波炉是如何加热的？
3.9　电饭锅可以用来烧水吗？
3.10　手机触屏是什么原理？
3.11　手机是如何定位的？
3.12　陀螺为什么不倒下？

# 第一章　有趣的数学

M——A——T——H





# 1.1　最早的学霸是谁？

——第一次数学危机

“学霸”这个词，现在就是指学习特别好的人。但是学霸这个词的本意并非如此，而是指那些在学术界利用自己的地位铲除异己，打压其他人的坏蛋，也叫“学术界的恶棍”。大家知道最早的学霸是谁吗？

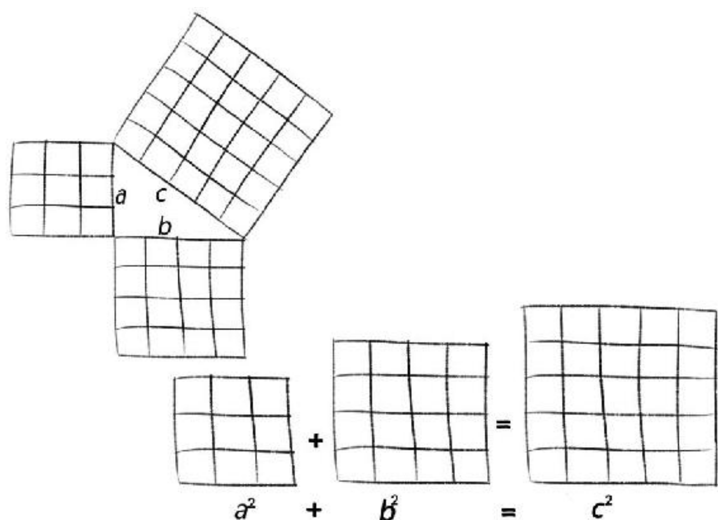



## 一、毕达哥拉斯：万物皆数

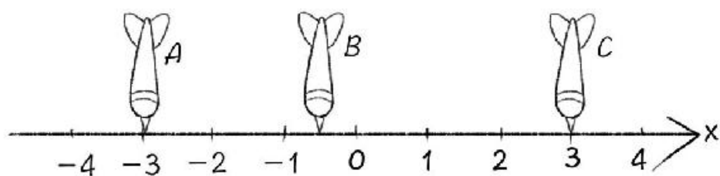
在公元前500年左右，也就是中国的春秋战国时期，地球上并没有多少国家。那个时期的希腊，以毕达哥拉斯为代表的毕达哥拉斯学派获得了丰硕的数学成果。



例如，他们提出了毕达哥拉斯定理，也就是我们熟知的勾股定理。这个定理告诉我们：一个直角三角形的两条直角边的平方和等于斜边的平方。




同时，毕达哥拉斯学派的观点是“万物皆数”，而且都是整数，他们认为宇宙的本质就是数，研究数学就是研究宇宙。




比如一根数轴，随便扔一支飞镖扎在数轴上，它可能会戳中一个整数点，比如$A$ 点（$x$ =-3）、$C$ 点（$x$ =3）；也可能戳中$B$ 点（$x$ =-0.5）。那么$B$ 点该如何写作整数呢？毕达哥拉斯说，这个点虽然不能写作一个整数，但是可以写作两个整数的比。

$-0.5=-\frac{1}{2}$。瞧，$B$ 点还是可以用整数表示！

# 二、“完美”的有理数

我们把可以写作两个整数之比的数叫作有理数，分数就是有理数，

而整数显然可以写作它自身与1的比，所以整数也是有理数。这样，毕达哥拉斯的观点可以概括为：数轴上任意一点都对应一个有理数。

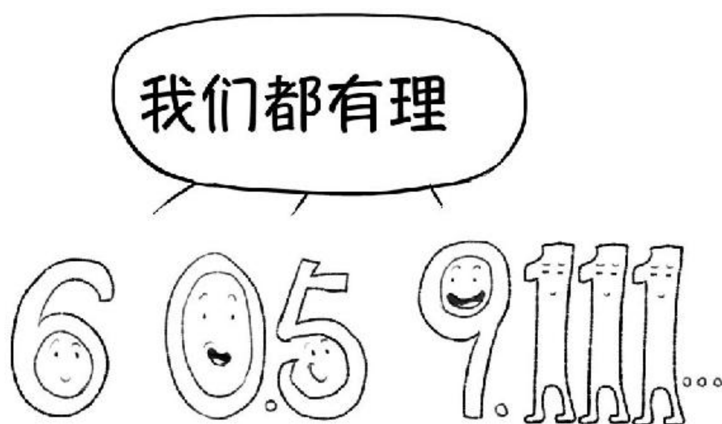
有理数可以分为三类：

第一类叫作整数，例如0、1、2、…

第二类叫作有限小数，它总可以写作一个分数的形式，例如：

$$3.5=\frac{7}{2}，\quad 3.8=\frac{38}{10}=\frac{19}{5}$$

第三类叫作无限循环小数，它也能写作一个分数的形式。例如 $0.333\cdots=\frac{1}{3}$



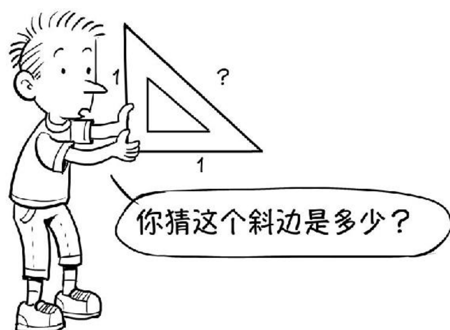

有同学可能要问，那0.343434…这个循环小数怎么写作分数呢？数学上把循环小数转化成分数的方法是首先数出循环节的位数，比如这里循环节是34，两位。然后再利用循环节除以循环节位数个9，也就是 $0.343434\cdots=\frac{34}{99}$。

看起来一切是那样的完美！

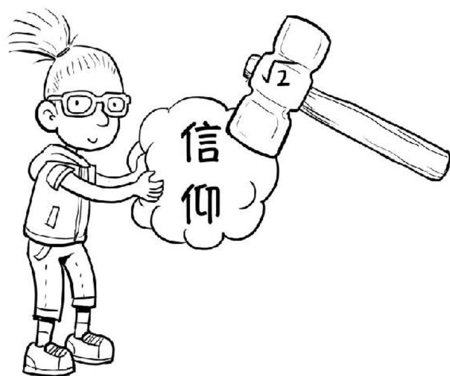
# 三、第一次数学危机

在毕达哥拉斯学派为自己的成就沾沾自喜时，有一天，学派内部一

个年轻学者希帕索斯提出了一个疑问。




“如果一个直角三角形的两个直角边都是1，那么斜边的长度如何表示成两个整数的比呢？”显而易见，根据勾股定理，斜边的长度是$\sqrt{2}$。但是$\sqrt{2}$如何才能表示成两个整数的比呢？希帕索斯为这个问题苦苦思索却没有答案，他只好求助于他的老师——毕达哥拉斯。




天真的希帕索斯以为毕达哥拉斯会给他答案，可谁知道，希帕索斯的问题居然动摇了毕达哥拉斯学派信仰的基础——万物皆是整数（或整数的比）。毕达哥拉斯实在无法解答这个问题，但是他又不想推翻自己已经建立的对数和宇宙的信仰。

最终，毕达哥拉斯选择了隐藏这个问题，他把可怜的希帕索斯扔进爱琴海里淹死了。希帕索斯成为历史上为探究真理而献身的人，而毕达哥拉斯则成了历史上第一位学术界的恶棍——学霸。

整个这件事，就被人们称为第一次数学危机。

# 四、危机是如何解决的？

为了解决这个问题，人们引入了无理数的概念：无理数就是无限不循环小数，无法表示成整数的比。例如圆周率π = 3.1415926…、自然对数的底e = 2.71828…、$\sqrt{2}$ 等，都是无理数。

现在我们知道，数轴上的点不是与有理数一一对应，而是与实数一一对应，而实数包含有理数与无理数两类。




在提出了实数的概念之后，人们又想，如果把-1开平方，结果等于多少呢？我们知道，任何实数的平方都是非负的，所以-1貌似不能开平方。但是人们受到当年希帕索斯的启发，认为这个数的存在还是有意义的，人们把这个数叫作虚数$i$，并且$i^2 = -1$。

后来，人们把实数和虚数又统称为“复数”。瞧，我们对数的认识越来越深刻了！

# 1.2　阿基里斯能追上乌龟吗？

——第二次数学危机

第二次数学危机是关于一个奇怪的数——“无穷小”的争论。这个争论的源头依然要追溯到古希腊时代。




## 一、芝诺悖论

约公元前495年，古希腊学霸毕达哥拉斯去世了。这时，一个5岁的孩子正在牙牙学语，他叫作芝诺。



芝诺也是古希腊数学家，他提出了一系列悖论以反驳时间和空间的连续性和变化问题，比如有一个悖论称为“阿基里斯永远追不上一只乌龟”。

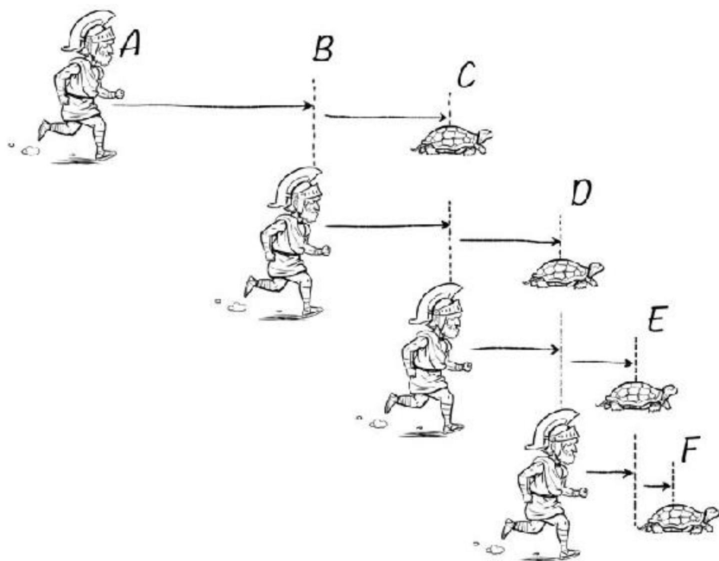
古希腊传说中有一位跑得最快的英雄阿基里斯，是希腊联军第一勇士，海洋女神忒提斯和英雄珀琉斯之子。阿基里斯出生后，忒提斯捏着他的脚踝将他浸泡在冥河斯堤克斯中，使他全身刀枪不入，唯有脚踝因被忒提斯手握着，没有浸到冥河水，这是他唯一的弱点。在特洛伊战争中，他被敌人射中脚踝而死。




有一天，阿基里斯遇到了一只乌龟。乌龟对阿基里斯说：“别看你跑得快，你永远也追不上我。”

阿基里斯问：“为什么呢？”

乌龟向他解释道：




开始比赛时，阿基里斯在后方$A$ 处，乌龟在前方$B$ 处，二者同时起跑。

阿基里斯要追上乌龟，首先要追上乌龟先跑的一段$AB$ ，但是在这段时间，乌龟也在向前跑，当阿基里斯到达$B$ 处时，乌龟已经跑到了$C$处，还没有追上。虽然此时$BC$ 的距离小于$AB$ 的距离。

阿基里斯会继续跑$BC$ 这一段，但是这段时间，乌龟也没闲着，它跑到了$D$ 处，虽然$CD$ 小于$BC$ ，但是阿基里斯还是没有追上乌龟。

以此类推，阿基里斯和乌龟之间的距离只能不断缩小，但是永远都不会变为零。所以，阿基里斯就永远追不上乌龟啦。

以上就是芝诺悖论。所谓悖论，一般是指同一个命题中有两个对立相反的结论。而芝诺对于阿基里斯追乌龟问题的解释不是推出对立的结论，而是完全违背常理，其实称为“诡辩”更加合适。

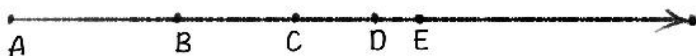



# 二、这个诡辩错在哪儿？

要推翻这个诡辩其实并不难。

芝诺将一个追及过程分割成无限多份：$AB$、$BC$、$CD$、$DE$、…并且认为，既然段数无穷多，那么累加起来的时间自然也是无穷长，所以追不上。但是实际上，由于阿基里斯速度快，乌龟速度慢，两人之间的距离会越来越短，时间也越来越短。无穷多个越来越小的时间之和，并不一定是无穷长。

为了更加清楚地解释这个问题，我们把追及过程画在一个数轴上，并且假设$AB$ 之间距离为$L$ 。为了方便起见，设阿基里斯的速度等于乌龟速度的两倍。




根据速度的二倍关系，相同时间内，阿基里斯运动的距离就是乌龟的两倍。所以阿基里斯走过$AB = L$ 时，乌龟走过的距离为$BC = \frac{L}{2}$ ；阿基里斯走过$BC$ 时，乌龟走过的距离为$CD = \frac{L}{4}$ ……显然，第$N$ 次追及的距离是$\frac{L}{2^{N-1}}$ ，如果$N$ 无限大，这个长度就无限接近于0，称为“无穷小”。

如果阿基里斯要追上乌龟，需要追及无限多段，将这无限多段距离求和：

$$S = L + \frac{L}{2} + \frac{L}{4} + \frac{L}{8} + \cdots$$

大家经过简单计算就会发现：项数越多，这个式子的结果就越接近$2L$。如果项数无穷多，阿基里斯跑过的距离就与$2L$相差无穷小。直到阿基里斯追上了乌龟，他跑过的总路程也不会超过$2L$。同样，如果阿基里斯跑过第一段距离$AB$的时间是$t$，那么无穷多段之后阿基里斯追上乌龟，需要的总时间不过是$2t$。




庄子曰："一尺之捶，日取其半，万世不竭。"说的就是一根一尺长的木棍，每天砍掉它的一半，无论经过多久，都砍不光。的确，我们可以把木棍分割成越来越短的无限多份，但是加起来依然是一根木棍那么长，这与芝诺悖论多么相似！

# 三、无穷小和导数：微积分的基础

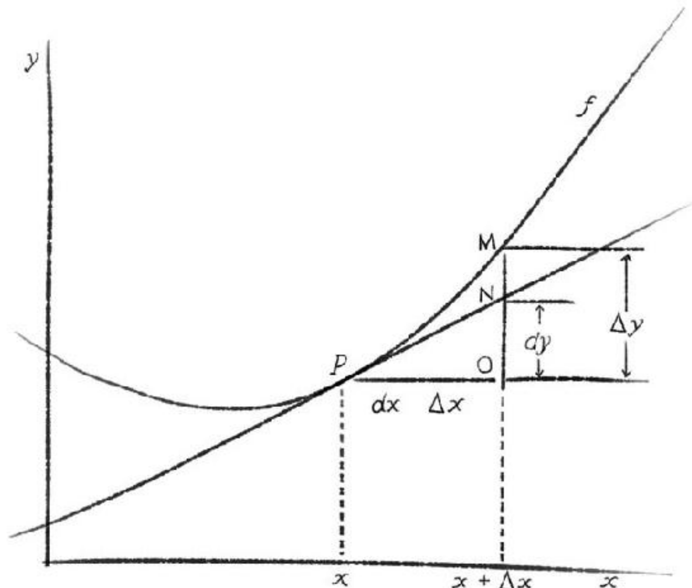
芝诺最早提出了无穷小的概念。只可惜，希腊文明衰落之后，欧洲的科学一直没有太大进步。直到文艺复兴时代来临，在牛顿等一大批科学家的带领下，科学才重新蓬勃发展起来。



我们都知道牛顿是伟大的物理学家，但他也是伟大的数学家。牛顿提出的牛顿二项式定理、牛顿二分法以及微积分，都是近代数学的辉煌成就。

微积分在物理上的应用非常广泛，人们利用它解决了很多复杂的问题。微积分中有一个概念：导数。

在一个函数图像上随便取两个点$P$和$M$，计算二者纵坐标差$\Delta y=y_2-y_1$与横坐标差$\Delta x=x_2-x_1$，那么$\frac{\Delta y}{\Delta x}$就称为“两个点连线的斜率”。




如果$M$ 点越来越接近$P$ 点，那么$PM$ 的连线就会变成过$P$ 点的切线，而这条切线的斜率就称为“$P$ 点的导数”。导数可以写作$f'(x)=\lim_{\Delta x\to 0}\frac{\Delta y}{\Delta x}$，这里lim就叫“极限”，$\Delta x$ 就称为“无穷小”。




# 四、第二次数学危机

本来一切看起来都很自然，但是英国大主教贝克莱首先对微积分发难：“无穷小到底是一个什么样的数？它是0吗？如果是0，为什么在求

导时可以做分母？如果不是0，又怎么能说刚才计算的是一个点的切线斜率，而不是两个点连线的斜率呢？”

从牛顿发明微积分到十九世纪二十年代，关于无穷小到底是怎样一个数，一直没有一个统一的认识。后来，经过阿贝尔、柯西、康托尔等人的努力，直到十九世纪七十年代，人们才确立起极限基本定理，使得无穷小有了一个合理的解释。从牛顿发明微积分开始计算，已经过去了300年。而从芝诺最早提出的无穷小概念，已经过去了2500年。

数学就是这样，为了一个看似简单的概念，可以争论几百年，甚至几千年。

# 1.3　我说的是假话

——第三次数学危机

“我说的是假话”看似平淡无奇，但是在数学上，却称得上是一个悖论。因为如果这句话是真的，按照字面意思，它就是一句假话；如果这句话是假的，那么就会得到和字面意思相反的结论：这是一句真话。悖论就是一个论述却可以得到两个互相矛盾的结论。




## 一、理发师悖论

英国数学家罗素提出了与之相似的著名悖论：理发师悖论。

从前有一个村子，村子里只有一名理发师。这个理发师有点怪，他的理发店门口立了一块牌子，写着下图中这段话。




也就是说，村子里的人分为两类，第一类人会给自己刮胡子，第二类人从不给自己刮胡子。而这名理发师不给第一类人刮胡子，只给第二类人刮胡子。

# 二、集合论的问题

有一天，一名顾客来到店里看到了这块牌子，他就问了理发师一个问题。



理发师突然发现自己非常尴尬。因为他如果回答给自己刮胡子，他就是第一类人，按照他的规矩，就不应该给自己刮胡子；如果他不给自己刮胡子，他就是第二类人，按照规矩，他又应该给自己刮胡子。

当然，这只是罗素悖论的通俗说法。罗素悖论是关于数学中集合论的一个矛盾而提出的。

什么是集合呢？所谓集合，是由某些确定的元素构成的整体。例如：

$A=\{1,2,3\}$是一个集合，里面有三个元素，分别是1、2、3；

$B=\{x\mid x$ 是偶数$\}$是一个集合，包含所有的偶数，有无限多个元素；

$C=\{x\mid x$ 是拖拉机$\}$是一个集合，任何一辆拖拉机都是这个集合的元素。

但是集合的元素必须是确定的。所以有些概念不能构成集合，例如“美女的集合”就是一种错误的说法，因为一个人美不美会因为其他人的感受而异，不具有确定性。

元素与集合的关系有“属于∈”和“不属于∉”两种，比如“1”这个元素，它是集合$A$的元素，但是不是集合$B$的元素，写作$1\in A$，$1\notin B$。

明白了这些，我们就可以讨论罗素悖论的数学表达了。罗素说：设集合$S$是所有不属于自身的集合构成的集合，即$S=\{x\mid x\notin S\}$。那么，$S$是否属于自身呢？

若$S$属于自身，那么$S$就不满足集合所规定的元素性质，它不应该属于自身$S$；

若$S$不属于自身，那么$S$就满足集合所规定的元素性质，它应该属于自身$S$。

这样一来，这个集合就得到了自相矛盾的结果，与理发师悖论如出一辙。

其实，罗素主要是一个哲学家、逻辑学家、教育学家和文学家，并且获得了诺贝尔文学奖。

# 三、第三次数学危机

罗素为什么要提出这个数学悖论呢？

二十世纪初，数学界笼罩在一片喜悦祥和的气氛之中。法国大数学家彭加莱在1900年的国际数学家大会上公开宣称：数学的严格性，现在看来可以说是实现了。他说这句话是有依据的，那就是德国数学家康托尔所创立的集合论。




2000多年以来，人类一直没有弄清楚无穷的概念。比如全体正整数1、2、3、4、…和全体正偶数2、4、6、8、…都是无穷多个，那么它们谁更多呢？

也许有人会说：那一定是正整数多啊！多了3、5、7、…这些奇数。但是实际上两个无穷大这样比较是不行的。

康托尔利用集合论向人类指出：如果两个集合中的元素可以建立一一对应的关系，那么这两个集合的元素个数就是一样多的。比如正整数集合就可以和正偶数集合建立一一对应关系：每个整数的两倍刚好对应一个偶数，即$x \in \{整数\}$，$y \in \{偶数\}$，$y = 2x$。所以正整数集合和正偶数集合元素个数是一样多的。

集合论为数学奠定了坚实的基础，许多概念不清的问题利用集合论

得到了完美的解释。数学家希尔伯特高度赞誉康托尔的集合论是“数学天才最优秀的作品”，是“人类纯粹智力活动的最高成就之一”。

但是，从集合论诞生的那一天起，针对集合论的诘难和各种悖论就没有停止过。尤其以1902年罗素悖论最为有名。数学家们只享受了集合论带来的短暂的祥和，就又陷入了一种无法解决的危机之中，因为集合论已经成为了现代数学的基础，渗透到数学的各个分支之中，因此集合论的这个悖论才会引发这么多的关注。“理发师悖论”就被称为第三次数学危机。

由于这个问题迟迟得不到解决，康托尔承受着巨大的精神压力，最终精神失常，死在了哈勒大学的精神病院里。时至今日，第三次数学危机依然没有完美解决。数学家们只是通过人为添加一些限制条件以回避悖论的出现。

无论如何，康托尔作为最伟大的数学家之一，会永远被人类铭记。

# 1.4　3.1415926...

——圆周率的计算

圆周率π是一个十分重要的数，也是一个很神奇的数。从古希腊时代开始，由于科学研究和工程技术的需要，圆周率的计算就一直没有停止过。直到今天，圆周率依然是检验计算机计算能力的方法之一。日本某个无聊的出版社居然出了一本一百万位的圆周率的书《円周率1000000桁表》，全书只有一个数字：π。




## 一、阿基米德：π≈3.14



公元前300年左右，古希腊数学家欧几里得在著作《几何原本》里将几何的基础简化成几条公理。其中一条公理是：过一点以某个长度为半径可以作一个圆。根据相似形可知：任何一个圆的周长与直径的比都是一个常数，这个常数被称为“圆周率π”。

如果使用一根软绳测量圆的周长，再除以圆的直径，只能得到圆周

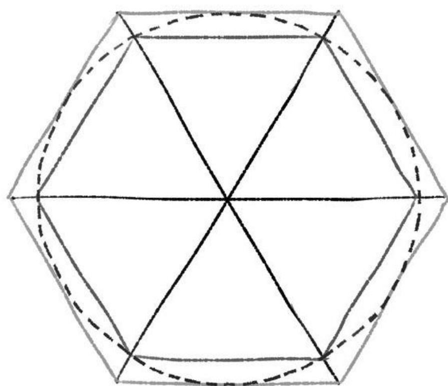
率大约等于3的结果，更加精确的结果只能依赖计算。

第一个把π计算到3.14的人是古希腊的阿基米德。

我们都知道阿基米德的名言：给我一个支点，我可以撬起地球。阿基米德是第一个发现杠杆原理和浮力定律的人，是一位物理学家。但是同时，他也是一位数学家。公元前212年，罗马士兵进攻叙拉古国，城破之后，阿基米德被罗马士兵杀死。传说他临死时被罗马士兵逼到一个海滩，还在海滩上画圆，并且对士兵说："你先不要杀我，我不能给后世留下一个不完善的几何问题。"




阿基米德计算圆周率的方法是双侧逼近：使用圆的内接正多边形和外切正多边形的周长来近似圆的周长。正多边形的边数越多，多边形周长就越接近圆的边长。



阿基米德最终计算到正96边形，并得出π≈3.14的结果。阿基米德死后，古希腊遭到罗马士兵的摧残，叙拉古国灭亡，古希腊文明衰落，西方圆周率的计算从此沉寂了一千多年。

## 二、刘徽和祖冲之：π≈3.1415926

阿基米德死后五百年，中国处于魏晋时期，著名数学家刘徽将圆周率推演到小数点的后四位。他在著作《九章算术注》中详细阐述了自己的计算方法。

刘徽的算法与阿基米德基本相同，但是刘徽提出了正$N$ 边形边长$L_N$与正$2N$ 边形边长$L_{2N}$ 的递推公式，并且计算到了圆的内接正3072边形，得到π的值大约是3.1416。



又过了两百年，中国数学家祖冲之横空出世。

祖冲之使用“缀术”将圆周率的值计算到小数点后第七位，指出：$3.1415926 < \pi < 3.1415927$，这个结果直到一千多年后才被西方超越。




但遗憾的是，“缀术”的计算方法已经失传。华罗庚等科学家认为，祖冲之的方法仍然是割圆法。但是如果要得到这个精度，需要分割到

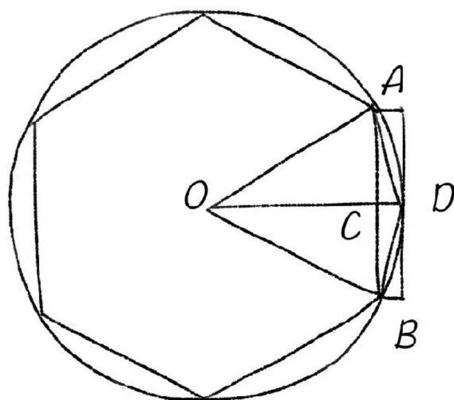
24576边形，从正六边形出发，还需要迭代刘徽的公式12次。而且在每次迭代的过程中，必须保证足够多的有效数字，否则就会影响到最后的结果。祖冲之通过什么神奇的方法保证了计算的准确性，至今仍是一个谜。

看到这里，也许有的同学已经跃跃欲试了，我们能不能仿照阿基米德和刘徽的方法自己计算一下圆周率呢？

# 三、割圆法到底是什么？

其实这个问题也没那么难，我们不妨也来简单推导一下：

首先作一个半径为1的圆。




设圆的内接正$N$边形的边长为$L_N$，如图中$AB$所示。

将正$N$边形变为正$2N$边形，边长$L_{2N}$，如图中$BD$所示。

在三角形$BCD$中列勾股定理：

$$BD=\sqrt{BC^2+CD^2}=\sqrt{\left(\frac{1}{2}AB\right)^2+(OD-OC)^2}$$

其中$AB=L_N$，$OD=1$，而$OC$又可以在三角形$OCB$中利用勾股定理计算：

$$OC=\sqrt{OB^2-BC^2}=\sqrt{1-(\frac{1}{2}L_N)^2}$$

根据以上计算可以得到递推式：$L_{2N}=\sqrt{2-\sqrt{4-L_N{}^2}}$ 。

当$N$ =6时，圆的内接正六边形边长刚好与圆的半径相等，$L_6$ = 1，用这个正多边形近似圆的周长，可以得到圆周率$\pi\approx\frac{6L_6}{2R}=3$ 。

当$N$ =12时，根据递推，得到$L_{12}=\sqrt{2-\sqrt{4-1}}=\sqrt{2-\sqrt{3}}$ ，用这个正多边形近似圆的周长，可以得到$\pi\approx\frac{12L_{12}}{2R}=3.1$ 。

按照这个方法一直计算下去，就可以得到更加精确的结果了。当然，时至今日，人们已经发明出各种各样计算π的方法。比如，欧拉就提出过使用级数方法计算π的值：$\frac{1}{1^2}+\frac{1}{2^2}+\frac{1}{3^2}+\cdots=\frac{\pi^2}{6}$ 。

这种方法要比使用割圆法快得多，也方便得多。

话说，你能背下来多少位的π呢？我能背下来小数点后22位，这是因为小时候看过的一个故事：

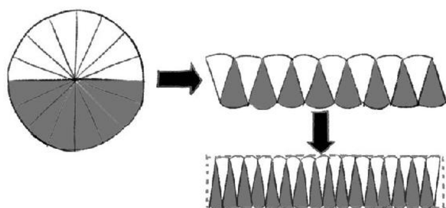
有位教书先生，整日里不务正业，就喜欢到山上找庙里的和尚喝酒。他每次临行前留给学生的作业都一样：背诵圆周率。开始的时候，每个学生都苦不堪言。后来，有一位聪明的学生灵机一动，想出妙法，把圆周率的内容与眼前的情景联系起来，编了一段顺口溜：

山巅一寺一壶酒（3.14159），尔乐苦煞吾（26535），把酒吃（897），酒杀尔（932），杀不死（384），乐尔乐（626）。

# 1.5　比萨饼中的数学

——微积分的基本思想

小学的时候我们就学过圆的面积公式$S=\pi r^2$，其中$S$是圆的面积，π是圆周率，$r$是圆的半径。大家还记得这个公式是怎么得到的吗？



## 一、圆的面积公式

首先，我们画一个圆，这个圆的半径为$r$，周长为$C$。我们知道，圆的周长与直径的比定义为圆周率，因此$C=2\pi r$，这个公式就是圆周率π的定义，是不需要推导的。




然后，我们把圆分割成许多个小扇形，就好像一个比萨饼分割成了很多小块。

我们把这些扇形比萨饼一正一反地拼在一起，这样就形成了一个接近于长方形的图形。

可以想象，如果圆分割得越细，拼好的图形就越接近长方形。如果圆分割成无限多份，那么拼起来就是一个严格的长方形了。而且，这个长方形的面积与圆的面积是相等的。我们要求圆的面积，只需求出这个长方形的面积就可以了。

显然，这个长方形的宽就是圆的半径$r$，而长方形的长是圆周长的一半$\frac{1}{2}C=\pi r$，根据长方形的面积公式“长方形面积 = 长 × 宽”，我们得到圆的面积公式：$S=r\times\pi r=\pi r^2$

其实，这个推导过程很简单，那就是先无限分割，再把这无限多份求和。分割就是微分，求和就是积分，这就是微积分的基本思想。

# 二、牛顿与莱布尼茨之争

大家知道微积分是谁发明的吗？

从古希腊时代开始，数学家们就已经利用微积分的思想处理问题了。比如阿基米德、刘徽等人，在处理与圆的相关问题时都用到了这种思想，但是那时微积分还没有成为一种理论体系。直到十七世纪，由于物理学中求解运动——天文、航海等问题越来越多，对微积分的需求变得越来越迫切。于是，英国著名数学家、物理学家牛顿和德国哲学家、数学家莱布尼茨分别发明了微积分。

1665年，牛顿从剑桥大学毕业，当时他22岁。他本来应该留校工作，但是英国突然爆发瘟疫，学校关闭了。于是牛顿只好回到家乡躲避瘟疫。在随后的两年里，牛顿遇到了他的苹果，发明了流数法，发现了色散，并提出了万有引力定律。



牛顿所谓的流数法，就是我们所说的微积分。但是牛顿当时并没有把它看得太重要，只是把它作为一种很小的数学工具，是自己研究物理问题时的副产品，所以并不急于把这种方法公之于众。

10年之后，莱布尼茨了解到牛顿的数学工作，与牛顿进行了短暂的通信。在1684年，莱布尼茨作为微积分发明的第一人，连续发表了两篇论文，正式提出了微积分的思想，这比牛顿提出的流数法几乎晚了20年。但是在论文中，莱布尼茨对他与牛顿之间通信的事只字未提。

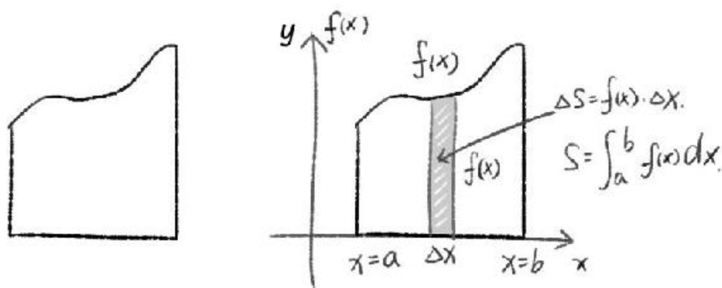
牛顿愤怒了。作为欧洲科学界的学术权威，牛顿通过英国皇家科学院公开指责莱布尼茨，并删除了巨著《自然哲学的数学原理》中有关莱布尼茨的部分。莱布尼茨也毫不示弱，对牛顿反唇相讥。两个科学巨匠的争论直到二人去世依然没有结果。我们今天谈到的微积分公式，还被称为“牛顿-莱布尼茨公式”。




他们在自己的著作中删除对手的名字，而后人却总是把他们的名字放在一起写。如果他们知道微积分公式现在的名字，又会作何感想呢？历史就是这么有趣。

# 三、微积分能干啥？

为了让大家更了解微积分及其应用，我们再来计算一个面积：有一块三条边为直线、一条边为曲线的木板，并且有两个直角。我们希望求出木板的面积。




为了求出这个面积，我们首先把木板放在一个坐标系内，底边与$x$轴重合。左右两个边分别对应着$x=a$和$x=b$两个位置，而顶边曲线满足函数$f(x)$。函数就是一种对应关系：每个$x$对应的纵坐标高度是$f(x)$。

如果我们把这个图形用与$y$轴平行的线进行无限分割，那么每一个

竖条都非常接近于一个长方形。长方形的宽是一小段横坐标$\Delta x$，高接近于$f(x)$，所以一个竖条的面积就是$f(x)\Delta x$。

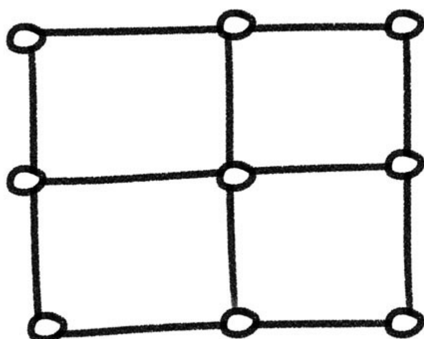
现在我们把无限多的小竖条求和，就是板子的面积，写作$S=\int_a^b f(x)dx$ 其中$a$叫作下限，$b$叫作上限，$f(x)$叫作被积函数，这个表达式就是积分，表示$f(x)$、$x=a$、$x=b$和$x$轴四条线围成的图形面积。

怎么样？虽然微积分的计算比较复杂，但是明白其中的原理还是十分简单的，对不对？

# 1.6　一笔能写出“田”字吗？

——欧拉与哥尼斯堡七桥问题

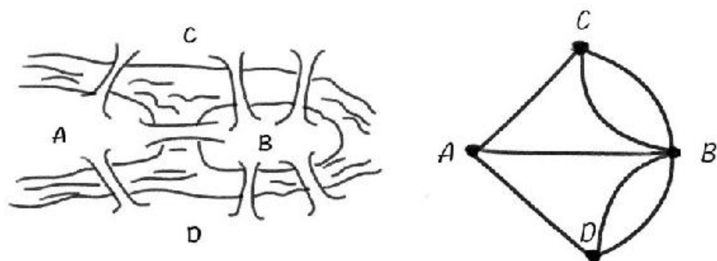
只用一笔，既不离开纸面，也不重复，能写出“田”字吗？这实际上是十八世纪一个经典的数学问题：哥尼斯堡七桥问题。



## 一、哥尼斯堡七桥问题

在普鲁士的哥尼斯堡（今俄罗斯加里宁格勒）有一个公园，公园里有七座桥，这七座桥将普雷格尔河中的两个岛屿与河岸连接起来。

1736年，当地居民举办了一项有意思的健身活动：一次走过七座桥，每座桥只能经过一次，而且起点与终点必须是同一座桥。




有许多人进行了尝试，但是都失败了。

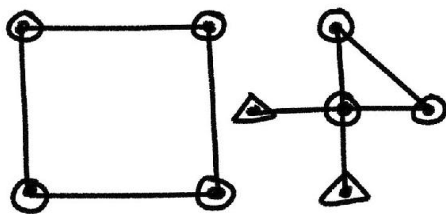
# 二、欧拉的“一笔画”




当时世界上最伟大的数学家欧拉刚好在这里，他敏锐地发现这里蕴藏着深刻的数学内涵，并把它称为“一笔画问题”。

欧拉把七座桥画作七条线段，并把问题转化为是否可以通过一笔将这个图形画出来。经过思考，欧拉认为这是不可能的。不仅如此，欧拉还得出了哪些图形可以一笔画，哪些不能一笔画的条件。

首先，欧拉把图形中的点分为两种：如果过该点的线段有偶数条，就称为“偶点”；如果过该点的线段有奇数条，就称为“奇点”。比如下面的图形中，圆圈的点就是偶点，三角形的点就是奇点。



欧拉指出：如果一个图形可以一笔画成，那么它的奇点个数一定是0个或者2个。

如果奇点个数是0个，那么起点和终点是同一个点，从图形中任何一点出发都可以一笔画，比如左上图就是这样。

如果奇点个数是2个，那么只能从一个奇点出发，画到另一个奇

点，才能将图形画出来，这就是右上图的情况。

理解这个问题其实并不难，因为：

如果一个点既不是起点，也不是终点，那么线段经过该点时必然会一进一出，线段成对出现，一定是偶点。

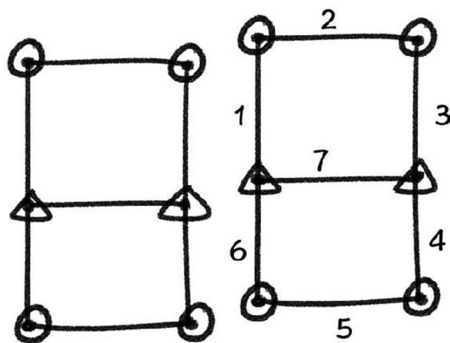
如果一个点既是起点，又是终点，那么该点有一条出发线段和一条结束线段，也是偶点。

如果这个点只是出发点，或者只是结束点，这样才可能是奇点。

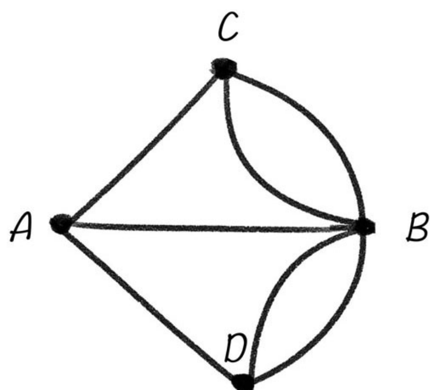
所以，如果从一点出发，一笔画回到这个点，图形中就不会有奇点；如果从一点出发，一笔画到另一点，图形中就会有两个奇点。

比如，我们来看看“日”是否能一笔画。

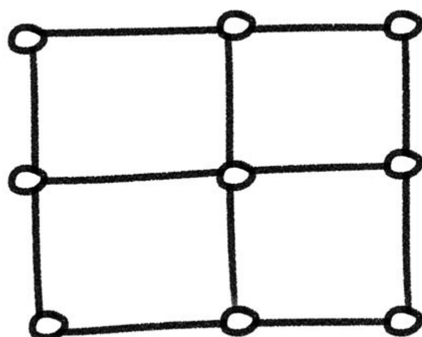
由于日字腰上的两个点有三条线段，因此是奇点，其余点都有两条线段，是偶点。因此日字可以一笔画，而且必须从腰上的一点出发到另一点结束。按照图中1、2、3、4、5、6、7的顺序，就能画出来了。




我们再来看看哥尼斯堡七桥问题。




在这个图形中，过$A$、$C$ 或$D$ 各有3条线段，是奇点；过$B$ 有5条线段，也是奇点。图中有4个奇点，因此上图是不能一笔画的。



说了这么多，读者是不是可以看看“田”字中有几个奇点？能不能一笔画成呢？

## 三、牛人欧拉




欧拉向圣彼得堡科学院提交《哥尼斯堡的七座桥》的论文时只有29岁，在解答问题的同时，他开创了数学的一个新的分支——图论与几何拓扑。

欧拉是一个天才，他在数学史上的地位就像牛顿在物理学史上的地位一样伟大，我们在研究数学时会经常看到欧拉公式、欧拉定理、欧拉函数。他13岁进入大学学习，16岁就获得了硕士学位。28岁时，由于生病，欧拉的右眼失明了。晚年时，左眼也失明了。但是就在双目失明的情况下，欧拉还凭借心算解决了许多数学问题。



他不光是数学史上里程碑式的人物，同时也是一位物理学家，为物理学的发展铺平了数学的道路。他一生中写出了886本书籍和论文。圣彼得堡科学院为了整理他的著作，足足忙了47年。

不要急，后面还会多次提到这位伟大的数学家。

# 1.7　1 + 1

——陈景润与哥德巴赫猜想

总有小朋友问我，科学家为什么要研究1 + 1 = 2这么弱智的问题？要讨论什么是“1 + 1”，要从十八世纪说起。



## 一、哥德巴赫和“1 + 1”




十八世纪初，也就是中国的清朝时期。俄罗斯伟大的君主彼得大帝修建了一座新城——圣彼得堡，并全面学习欧洲。

彼得大帝从欧洲引进了一批科学家来建设新的国家，其中就有德国

数学家哥德巴赫。哥德巴赫最初是一名中学教师，后来在圣彼得堡担任圣彼得堡帝国科学院教授，1728年开始担任彼得二世（彼得大帝的孙子）的宫廷教师。

哥德巴赫在研究中发现：很多偶数都可以分解成两个质数的和。

什么是质数呢？质数也被称为“素数”，只有1和它本身两个约数。比如，2、3、5、7、11、13、17等都是质数，因为除了1和它本身外，这些数都没有其他约数。

与之对应的另外一种数是合数：除了1和它本身，还有其他约数。比如，6是合数，因为它有约数1、2、3、6；8是合数，因为它有约数1、2、4、8；9是合数，因为它有约数1、3、9。




哥德巴赫的猜想就是：任何一个大偶数（$x \geq 4$）都可以被分解成两个质数的和。

比如，4 = 2 + 2，6 = 3 + 3，8 = 3 + 5，10 = 3 + 7。

是不是所有偶数都能这样呢？这就构成了一个猜想，并被称为“1 + 1”。

哥德巴赫无法证明这个猜想，于是写信求助著名数学家欧拉。欧拉是堪称“优秀本秀”的科学家，到目前为止，还没有哪位数学家的成就能超过欧拉。但欧拉这么厉害的人也无法解答这个问题。




于是这个问题流传下来，并困扰了数学界200多年。二十世纪，人们对这个问题展开围攻。

1920年，挪威的布朗证明了“9 + 9”。

1924年，德国的拉特马赫证明了“7 + 7”。

1932年，英国的埃斯特曼证明了“6 + 6”。

1937年，意大利的蕾西证明了“5 + 7”。

1938年，苏联的布赫夕太勃证明了“5 + 5”。

1940年，苏联的布赫夕太勃证明了“4 + 4”。

1956年，中国的王元证明了“3 + 4”。

1962年，中国的潘承洞证明了“1 + 5”。

1962年，中国的王元证明了“1 + 4”。

1965年，苏联的布赫夕太勃证明了“1 + 3”。

什么是“1 + 3”呢？就是说一个大偶数一定可以分解成一个质数及不超过三个质数乘积之和的形式，即证明了对于任何一个大偶数$x$ ，一定

可以分解成$x = a + b$或$x = a + bc$或$x = a + bcd$的形式，其中，$a$、$b$、$c$、$d$都是质数。

## 二、陈景润和“1 + 2”



那么我国数学家陈景润又干了些什么呢？

陈景润是我国著名数学家。他的中学数学老师是国立清华大学航空系的主任。他的老师上课时喜欢讲一些科学故事，比如哥德巴赫猜想。老师说，数学是科学的王后，而数论是王后的王冠，哥德巴赫猜想就是王冠上一颗璀璨的明珠。

陈景润对这颗明珠非常感兴趣，就一直致力于研究这个问题。后来他去了厦门大学读书，毕业后被分配到北京一所中学当数学老师。

陈景润不善于与人交流，讲课讲得很差，跟学生的关系也不好，还经常生病，有人说他是高分低能。但是厦门大学校长慧眼识珠，认定陈景润是厦大最优秀的毕业生，于是干脆把陈景润调回厦大工作。

陈景润回到厦大后专心研究数学，并把他的研究成果寄给了北京的华罗庚。华罗庚当时已是享誉全球的数学家，一眼看中陈景润，便把他调到中科院数学所担任研究员。

陈景润回到北京后，还是不与人交流。而且当时正在“文革”期间，

学术环境很不好。但是就在这样艰苦的条件下，他却用几麻袋的演算纸证明了“1 + 2”。

任何一个大偶数都可以分解成一个质数及不超过两个质数的乘积之和。即证明了对于一个大偶数$x$ ，可以被分解成$x = a + b$ 或$x = a + bc$ ，其中，$a$、$b$、$c$ 都是质数。

从“1 + 3”到“1 + 2”，看似是一小步，但实际是一个很大的成就，被称为“陈氏定理”，获得了国际公认，也受到了周恩来总理和毛主席的肯定。




遗憾的是，陈景润依然没有证明“1 + 1”，虽然看起来他距离“1 + 1”只有一步之遥，但直到今天，“1 + 1”仍然是一个谜。我们也正期待着新的科学巨匠的出现。

所以，“1 + 1”的含义是一个质数加一个质数，而不是“1 + 1”为什么等于2。

# 1.8　最厉害的民科是谁？

——费马大定理

“民科”最初是民间科学爱好者的简称，意思是“爱好科学，但是并不在科学体系内工作的人”。这回带大家了解一下被称为“民科之王”的法国律师——费马。




## 一、费马的一个猜想：费马数

费马是十七世纪的法国律师，业余时间研读了很多数学著作，经常会提出自己的猜想，而且他的思想与眼光一点也不比专业数学家差，对解析几何、微积分、数论、物理学都有重大贡献，是那个时代法国最伟大的数学家。

他经常会去图书馆阅读数学书籍，而且会在书的空白处写下自己的猜想，由此诞生了许多数学史上困扰人们的难题。有的难题困扰了世界几十年，甚至有的困扰了几百年。

例如，费马曾经提出猜想：$2^{2^n}+1$ 对于所有的正整数$n$ 都得到一个质数。

$n=0$，$2^{2^0}+1=2^1+1=3$ 是质数。

$n=1$，$2^{2^1}+1=2^2+1=5$ 是质数。

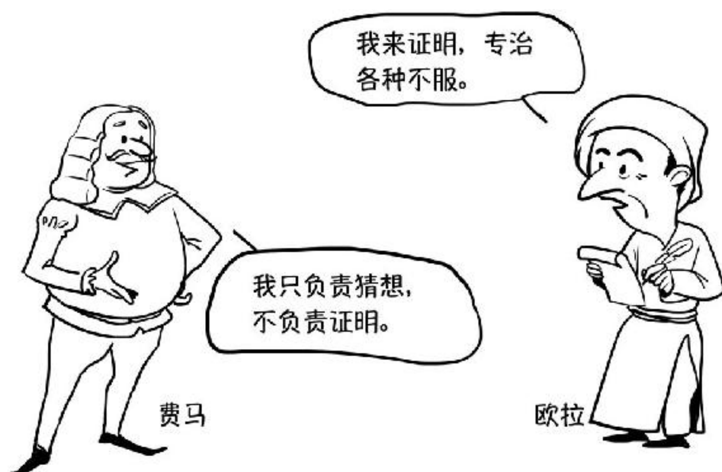
$n=2$，$2^{2^2}+1=2^4+1=17$ 是质数。

$n=3$，$2^{2^3}+1=2^8+1=257$ 是质数。

$n=4$，$2^{2^4}+1=2^{16}+1=65537$ 是质数。

1640年，费马提出这个猜想，使得无数数学家绞尽脑汁思索了90多年，直到“优秀本秀”的数学家欧拉出现之后，这个问题才得以解决。

欧拉在1732年出：费马这个猜想是错误的，因为$n=5$，$2^{2^5}+1=2^{32}+1=4294967297=641\times6700417$不是一个质数。




有人会问，再算一个数就可以了，为什么需要花90多年的时间呢？这是因为大数的质因数分解是非常困难的，人们还没有找到一种方法可以判断一个很大的数是不是质数，只能通过一个一个约数进行尝试。我们目前的密码学也是基于这一原则。

不过，欧拉证明了$n=5$不是质数之后，人们又计算了几个费马数，发现从$n=5$开始，直到$n=11$，都不是质数。比如$n=11$时，这个数是：

$2^{2048}+1=$

32,317,006,071,311,007,300,714,876,688,669,951,960,444,102,669,71

它等于

319,489×974,849×167,988,556,341,760,475,137×3,560,841,906,445

于是，人们又猜想，是不是从$n=5$开始，费马数都是合数呢？虽然我们已经有了计算机，但是这个问题到现在还没有解决。不仅如此，就连$n=12$这个数如何进行质因数分解都还不知道，它还有一个1187位的因子没有被分解。因为费马数实在是太大了。

# 二、费马的另一个猜想：费马大定理

相比于费马数，费马还有一个更著名的猜想，现在称为“费马大定理”或“费马最后定理”。

“当整数$n>2$时，关于$x$，$y$，$z$的方程$x^n+y^n=z^n$没有正整数解。”

$n=1$，这个方程变为$x+y=z$，显然有无穷多解；

$n=2$，这个方程变为$x^2+y^2=z^2$是勾股数，也是无穷多组解；

那么，$n=3$、4、5、…时，有没有正整数解呢？

费马管杀不管埋，1637年提出猜想后自称已经证明，但空白太小就不写了。150年后，神一般的数学家欧拉也只证明了$n=3$时没有正整数解的情况。

这个问题困扰了数学界350年，在这350年中，世界上一流的数学家如欧拉、高斯、刘维尔、柯西等人都参与过猜想的证明。这个定理一次次被人们宣布证明完毕，又一次次被证明推导过程是有问题的。但是在人类向这个猜想发起挑战的过程中，产生了许多数学成果，很多数学分支都由此而诞生。

还有一件有趣的事情。1908年，德国实业家沃尔夫斯凯尔临终时设

立了沃尔夫斯凯尔奖：凡是能在他死后100年内证明费马大定理的人，可以得到10万马克的奖励。他之所以设立这个奖项，是因为他年轻时为情所困，有一天决心在午夜自杀。但是晚上偶然看到了费马猜想，就饶有兴趣地开始证明，不知不觉天都亮了，他发现数学这么有意思，也不想自杀的事了。后来成了企业家之后，念念不忘费马的救命之恩，所以就设立了沃尔夫奖。




1955年，日本数学家谷山丰提出了谷山丰猜想，是一个关于椭圆曲线的问题。人们发现这个曲线似乎与费马猜想有着密切的联系。最终在1995年，英国数学家安德鲁·怀尔斯证明了费马猜想，费马猜想由此更名为费马大定理。怀尔斯把证明发表在《数学年刊》第142卷，实际占满了全卷，共五章，130页，题目为《模形椭圆曲线和费马大定理》。1996年，怀尔斯获得沃尔夫奖。1998年，怀尔斯获得数学界的诺贝尔奖——菲尔兹奖。

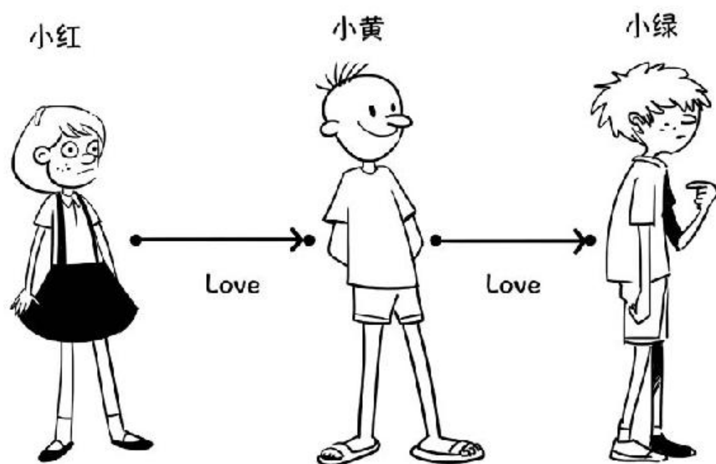
费马一生从没接受过专业的数学教育，却成为了十七世纪法国最伟大的数学家，不仅如此，他还让后代无数数学家为之痴狂。让我们记住这个超级“民科”的名字：费马。

# 1.9　怎样传小纸条？

——密码学的基本原理

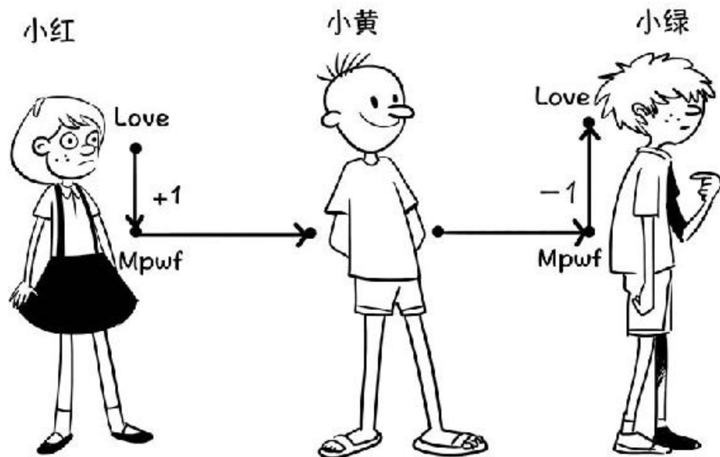
大家上学的时候是不是都传过小纸条？传小纸条的时候怎么才能不被人偷看呢？

假如有一个人名字叫作小红，她想把一句话“Love”传给小绿，但是由于距离远，必须通过一个人小黄。在这个过程中，小黄可能会偷看纸条内容。怎么才能防止小黄偷看呢？小红决定对信息进行加密。




# 一、对称加密

放学的时候，小红偷偷告诉小绿：“以后我给你写的话，都会往后推一个字母，比如L就变成了M，o就变成了p，v就变成了w，e就变成了f。你收到纸条之后，把纸条内容减去一个字母，就知道我想说什么了。这样一来，就算小黄偷看了我们纸条上的内容，也不知道我们要说什么。”




以上的过程就是密码学中最基本的加密算法：对称加密。也就是说，我们把明文（Love）按照一定的密钥（-1）加密成密文（Mpwf），对方接收后，再利用同样的密钥（-1）进行解密，就再次得到明文（Love）。

但是这种加密方法面临很多的问题。比如，小黄虽然不知道密钥是什么，但是他可以一次次地用各种方法尝试密钥。比如，在英文中26个字母出现的频率是不同的。只要截获了大量的密文，就可以利用频率法猜出密钥，从而破解密码。

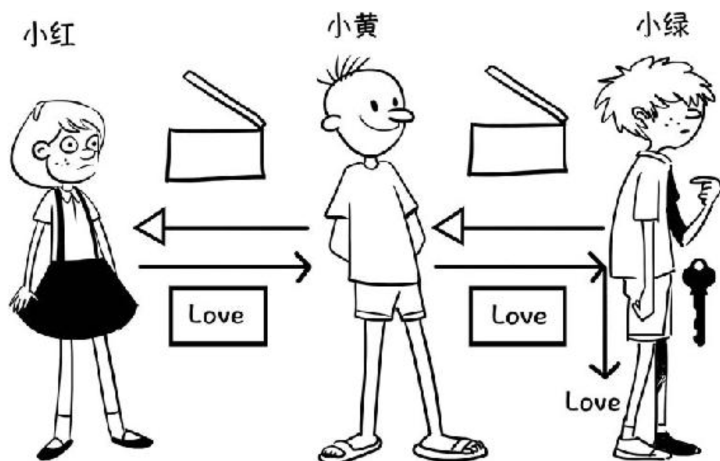
为此，小红和小绿只好不停地更换密钥，每天晚上放学都要商量一下第二天的密钥是什么。但是万一哪天两人放学没有商量好，或者商量密钥的时候被小黄偷听到了，那他们的密码就有可能被破译。商量密钥的过程就称为“密钥分发”，而密钥分发是对称加密算法最大的风险。

那怎么办呢？

# 二、非对称加密

两人想出了一种新的方法：首先小绿拿着一个没有锁上的空盒子，这个盒子只要一扣，就可以锁上。他让小黄把箱子传给小红，然后小红把小纸条放进盒子里，把盒子扣上。再通过小黄把盒子传给小绿。盒子的钥匙只有小绿有，小绿拿到盒子之后，用钥匙打开，就可以拿到小纸

条了。




这种方式就是现代更加保密的加密方式：非对称加密。也就是加密过程（锁箱子）方法是公开的，而解密过程使用的密钥（钥匙）是不公开的，而且加密过程的密钥（锁）和解密过程的密钥（钥匙）并不相同。小黄可以截获箱子，也知道加密方法，但是由于没有钥匙，他无法打开箱子，所以就不知道信息内容是什么了。

有同学可能要问，那么小黄不能通过一次次尝试找出钥匙吗？这就取决于这把锁是否足够复杂了。

通过前面两小节的内容，我们已经知道了大数的质因数分解非常困难，在密码学中也利用了这一点设计加密和解密算法。这种算法除了穷举，还没有找到更快的计算方式，而穷举所花费的时间非常长，从而保证了密码的安全。而且，小绿可以不停地更换锁头和钥匙，这个过程无须与小红进行沟通，也就解决了密钥分发的问题。

# 三、RSA算法

那么，具体的过程是怎么实现的呢？我们来介绍一种经典的加密算法：RSA算法。

RSA算法是1977年麻省理工学院的三名数学家罗纳德、萨莫尔、阿

德曼一起提出的，RSA就是他们三个名字的首字母组成的。这种加密算法基于欧拉定理等数学工具。具体算法是：

假如小红要把一个数字$m$ 传输给小绿，那么过程是这样的：

小绿首先找出两个质数$p$ 和$q$ ，计算它们的乘积$n$ ，并把这个乘积传递给小红，在传输过程中，小红可以利用$n$ 对传输的内容Love进行加密，并把加密后的结果返回给小绿。小绿拿到密文后，需要利用原来的两个质数$p$ 和$q$ 才能够进行解密。在这里$n$ 相当于箱子（公钥），$p$ 和$q$ 相当于钥匙（私钥）。

在这个过程中，小黄可以截获大数$n$ 和密文，但如果要解密，必须使用私钥$p$ 和$q$ ，要知道$p$ 和$q$ ，就必须对$n$ 进行质因数分解。

根据我们之前所说的，大数的质因数分解非常困难，计算一个费马数都花了90多年的时间。虽然现在有了计算机，而且计算能力飞速提高，但是报道过的曾被破解的RSA算法中$n$ 最大只有768位二进制数，而现在所使用的RSA算法大数$n$ 普遍有1024、2048或4096位二进制数，这么大的数字在有限的时间内，计算机也无法进行质因数分解，于是就保证了密码的安全性。

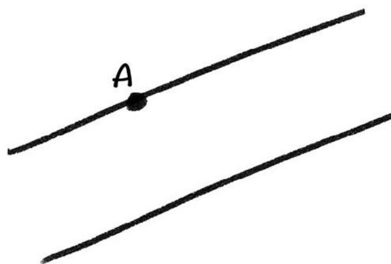



不过，根据科学家的研究，如果量子计算机被发明出来，大数的质因数分解时间就会大大缩短，那么传统密码就会面临风险。不过大家不用担心，到时候，科学家们会想出更好的方法进行加密的。

# 1.10　平行线存在吗？

——欧式几何与非欧几何

什么是平行线？许多同学都会说：太简单了，就是两条不相交的直线。而且我们在初中就学习过，过直线外一点，只能做一条已知直线的平行线。




其实，这个问题并没有那么简单，人们认清平行线的问题花了2000多年。

## 一、欧几里得几何

公元前300年，古希腊数学家欧几里得写了一本著作《几何原本》。



这本书共15卷，研究深入透彻，两千多年以来一直是数学几何部分的主要教材，被翻译成多种文字在世界上流传。直到今天，数学家们仍然要通过借助这本书中平面几何的点、线、面和立体模型来开展数学研究。

在《几何原本》中，欧几里得提出了五个基本假设，即所谓“公设”，公设是不可证明的。

欧几里得的五个公设是：

1．任意两点确定一条直线。

2．任意线段能延长成一条直线。

3．以一点为圆心、一个线段为半径可以作一个圆。

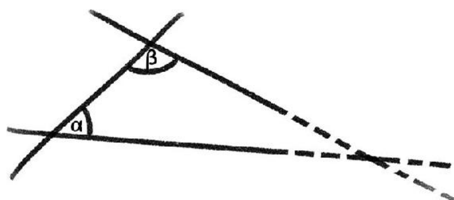
4．所有直角都相等。

5．过直线外一点有且只有一条直线与已知直线平行。

欧几里得从这五个公设出发，推导出一系列定理。这五个基本假设连同推导出的定理，被称为“欧式几何”，也就是我们中学时学习的几何。比如一个典型结论就是：三角形的内角和是180度。

# 二、罗巴切夫斯基几何

虽然几何研究取得了许多令人满意的成果，但是人们发现：第五公设的表述比较复杂。现在我们说的是简化版本，欧几里得最初的表述是：若两条直线都与第三条直线相交，并且在同一边的内角之和小于两个直角和，则这两条直线在这一边必定相交。



一些数学家怀疑这并不是公设，而是可以通过前四个公设推导出来的定理。于是，在很长一段时间，很多数学家都试图攻克这一难题，从前四个公设推导出第五公设，但大多无功而返。

第一个获得突破的人是俄罗斯数学家罗巴切夫斯基。

罗巴切夫斯基的父亲也是数学家，曾为了证明第五公设而耗尽一生。当他得知自己的儿子也开始研究第五公设的证明时，写信给他儿子说："千万不要研究这个问题。我几乎研究了所有的方法，最后都失败了，我不希望你也陷入这个泥潭。"

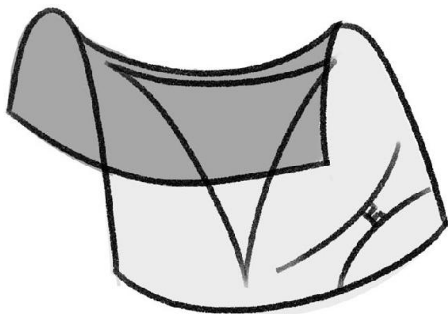
然而，罗巴切夫斯基并没有听从父亲的建议。他采用了一种与前人不同的方法：前人都是研究如何从前四个公设推出第五公设，而罗巴切夫斯基却反其道而行之，将第五公设修改为"过直线外一点至少有两条直线与已知直线平行"。

那么，假如第五公设可以证明，修改后，它必然与前四个公设相互矛盾，于是通过前四个公设以及修改后的第五公设通过演绎方法推导平面几何定理，一定能找到这个矛盾，然后就可以顺藤摸瓜证明第五公设了。

按照这个思路，罗巴切夫斯基用欧几里得的前四个公设与修改后的第五公设推导了平面几何中的所有定理，而且没有发现矛盾。他终于明白：第五公设的确是公设，不可以通过前四个公设证明。

既然第五公设不可证明，是一种假设，那么我们也可以更改这种假设。于是，罗巴切夫斯基将第五公设改为：过直线外一点有多条直线与已知直线平行，创立了自己的几何：罗氏几何。

罗氏几何中很多规律与欧式几何不同，比如最典型的三角形的内角和。在罗氏几何中，三角形的内角和不是180度，而是小于180度，具体的数值与三角形的面积有关：三角形面积越大，内角和越小。如果三角形面积无限大，内角和甚至可以为零。我们可以想象在双曲面上画三角形，内角和就小于180度，所以罗氏几何也叫作双曲几何。



1826年，34岁的罗巴切夫斯基在喀山大学物理数学系学术会议上，宣读了他的第一篇关于非欧几何的论文。但是论文一提出，就受到传统数学家们的嘲讽，人们对非欧几何的评价是“莫名其妙”。后来，罗巴切夫斯基被推选为喀山大学校长，即便如此，数学界对他的成就始终没有认可。

其实，在罗巴切夫斯基提出非欧几何理论时，当时世界上最顶尖的数学家、与欧拉齐名的数学之王、德国数学家高斯早就有了这种思想的萌芽。但是，高斯由于害怕新几何会激起学术界的不满和社会的反对，进而影响他的尊严和荣誉，所以生前一直没敢把自己的这一重大发现公之于世。当他看到罗巴切夫斯基的工作后，私下里向朋友说罗巴切夫斯基是俄罗斯最优秀的数学家，并决心学习俄语来了解罗巴切夫斯基的著作。但是公开场合却从没有对非欧几何进行任何支持。



罗巴切夫斯基在抑郁中走完了自己的一生。但是历史最终给了他公正，他死后人们才认识到非欧几何的巨大意义。1893年，在喀山大学竖立起了世界上第一个为数学家而雕塑的塑像。这位数学家就是俄罗斯的

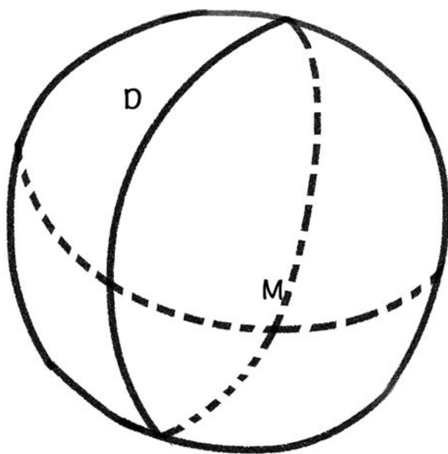
伟大学者、非欧几何的重要创始人——罗巴切夫斯基。

# 三、黎曼几何



那么，既然罗巴切夫斯基可以把第五公设中平行线条数改为多条，我们也可以改为一条都没有。于是，德国数学家黎曼将第五公设修改为：过直线外一点没有任何一条直线与已知直线平行，就创立了黎曼几何。

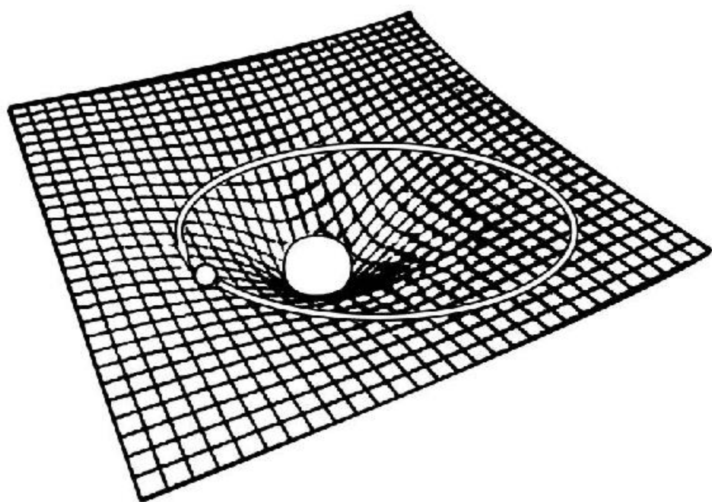
黎曼几何类似于在椭球上的几何，所以也称为“椭圆几何”。




在黎曼几何中，直线可以无限延长，但是总长度是有限的，而且不存在平行线的概念。也许有同学会问，地球的两条纬线难道不是平行的吗？简单来讲，在球面上，只有大圆（圆心在球心的圆）才能称为“直

线”，而任意的两个大圆都必然是相交的。

黎曼几何在爱因斯坦广义相对论中有很大的作用，这是因为空间本身并不是平直的，而是弯曲的。传说爱因斯坦在研究广义相对论时遇到了很大的数学困难，直到他发现黎曼几何这个有力的工具，才顺利地用数学表达了自己的思想。



数学就是这样，从假设与逻辑出发，演绎出许多结论。数学家们在研究数学时也许并不知道在生活中有什么应用，但正是数学家开创性的工作，才让其他学科的科学家更加方便地解释这个世界。



# 1.11 四维空间是什么样的？

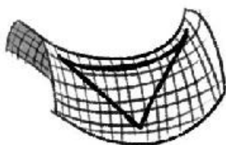
——欧式高维空间

我们经常在科幻电影或者科幻小说里听到四维空间，有人说，四维空间就是三维空间加上时间轴，这种说法是错误的。三维空间加时间轴是在相对论中使用的闵可夫斯基四维时空，而不是传统的四维空间。

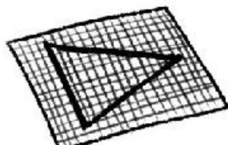
上一节我们说到了几何分为欧几里得式几何和非欧几里得几何，它们的区别在于第五公设的不同。有时候我们也称欧几里得几何空间为平直空间，而非欧几里得空间称为“弯曲空间”。



正曲率

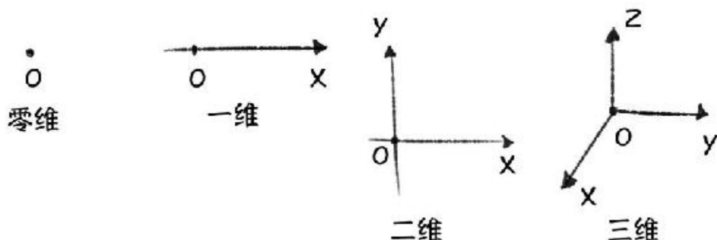


负曲率



平面曲率

我们这一次就来研究一下欧几里得几何中的高维空间是什么样子的。三维以下的欧式空间是比较好理解的：零维就是一个点、一维就是一条线、二维就是一个平面、三维就是一个立体。我们的想象力就到此为止了，因为我们生活的空间就是三维的。




但是，维度是一个数学概念，数学可以解释宇宙，但是数学本身并不需要为宇宙负责，也就是说，数学可以突破宇宙维度的限制。我们可以按照前面低维空间的规律向上推演，这样就可以得到高维空间的性质了。

# 一、高维立方体

首先，我们会发现：$N$ 维度空间就是过空间中任意一点可以作$N$ 条相互垂直的直线的空间。

比如：零维空间就是一个点，不能作出相互垂直的直线。

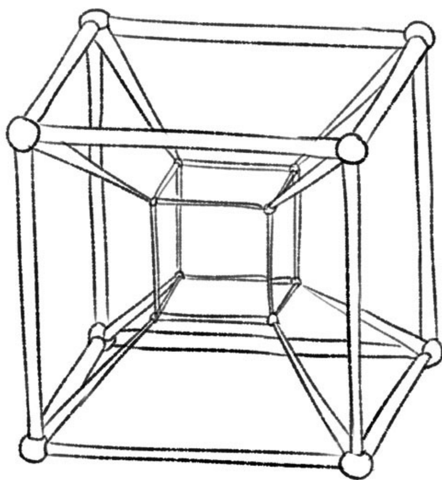
一维空间是一条线，过其中任意一点可以作出一条直线。

二维空间是一个面，过其中任意一点可以作出两条相互垂直的直线。

三维空间是立体，过空间中任意一点可以作出三条相互垂直的直线。

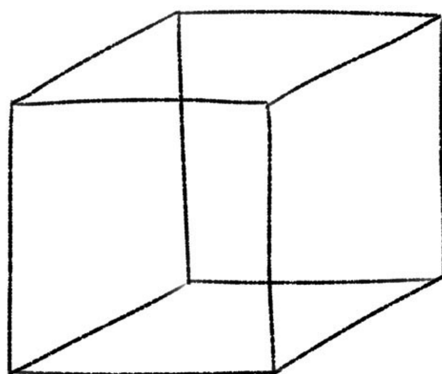
以此类推，$N$ 维空间就是过空间中任意一点可以作出$N$ 条相互垂直的直线的空间。

以四维空间为例，过其中任意一点可以作出四条相互垂直的直线。我们可以想象成两个三维立方体可以在这个空间中进行点点连接，造成每过一顶点都有四条直线相互垂直。

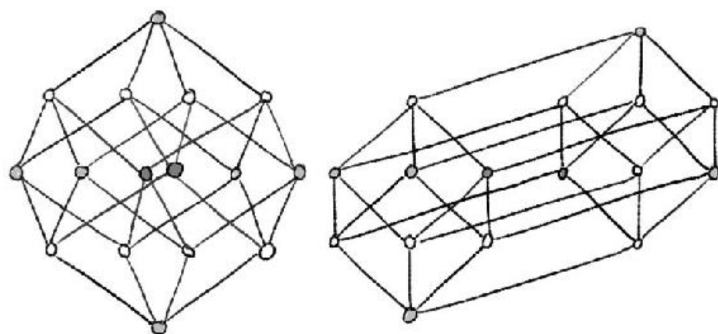


也许有同学说：不对啊！过每个点都有四条线，我看出来了，但是这四条线并不是相互垂直的啊！

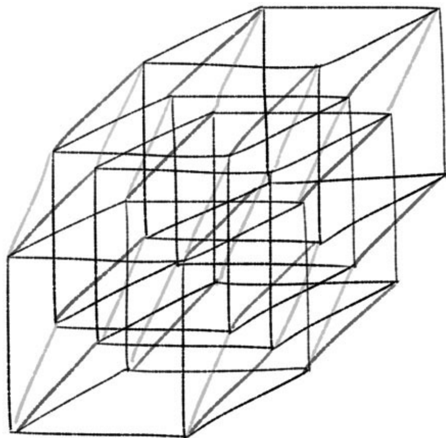
这是因为我们生活在三维空间中，没办法画出四维图形，只能画出四维图形在三维空间中的投影。这就好像我们在纸上画一个立方体，立方体的每条边看起来也不是相互垂直的一样。只不过我们生活在三维空间中，我们可以想象出这个立方体，所以在脑子中我们认为三条边是垂直的。



但是，四维空间我们没见过，所以不太好想象四维立方体中过一点的四个边是如何相互垂直的。而且，选取的投影面不同，四维立方体在三维空间中的投影图形长得也不一样。比如下面两种画法，就是四维立方体在三维空间的不同投影图。



同理，五维空间就是过其中任意一点可以作出五条相互垂直的直线的空间。图形画起来就更复杂啦。



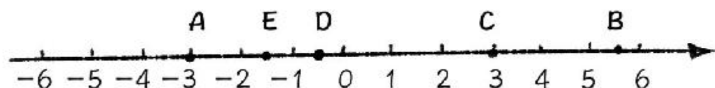
# 二、高维空间中的点

我们如何才能确定空间中的一个点呢？这就需要使用解析几何的知识了。解析几何是由法国学者笛卡尔开拓的，是一种利用代数方法研究几何问题的数学分支。

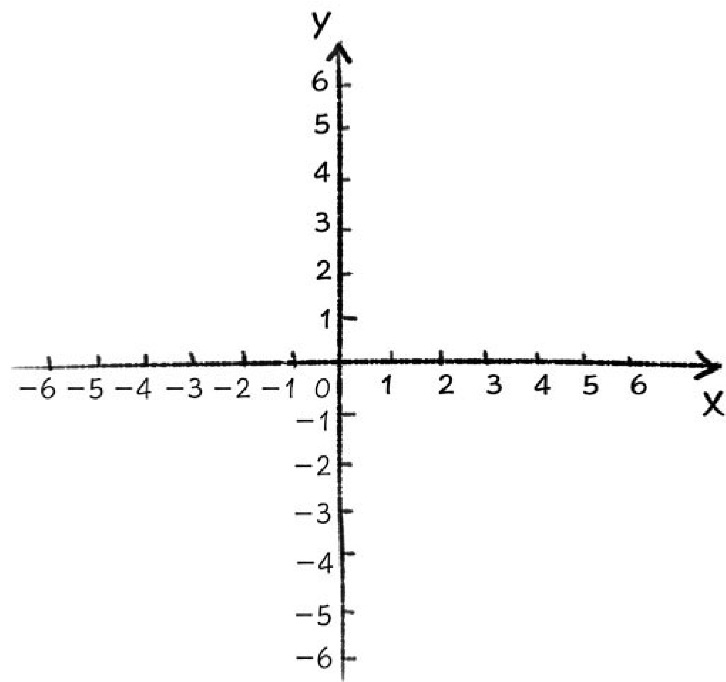
利用解析几何，我们可以在空间中建立一个坐标系，坐标系就是由相互垂直的数轴构成的。这样一来，空间中的点都可以用一个坐标表示。那么，$N$ 维空间中的坐标点包含有$N$ 个数字。

比如：

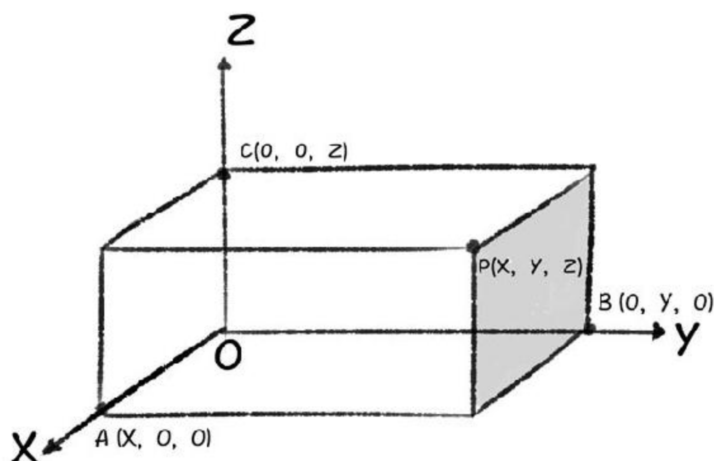
一维空间坐标系就是一根数轴，每个点的坐标都是一个数字；




二维空间坐标系称为“平面直角坐标系”，是两根相互垂直的数轴，每个点的坐标是一对数字（$x$ ，$y$ ），分别称为横坐标和纵坐标。




三维空间坐标系是三条彼此垂直的数轴，空间中每一个点的坐标都是三个数字（$x$ ，$y$ ，$z$ ），分别表示该点在三个数轴上对应的坐标。



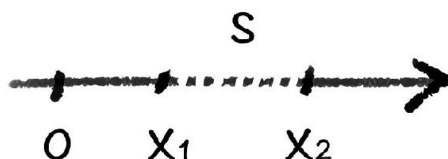

这样一来，$N$ 维空间中的点就需要一个$N$ 维坐标表示，也就是说，需要$N$ 个数字合并在一起，才能表示出空间中的一个点。如果是四维空间中的一个点，就需要四个坐标（$x$ ，$y$ ，$z$ ，$t$ ）才能把它表示出来。

# 三、高维空间的距离

而且，我们还可以定义出“空间距离”这个量。在低维空间中，两个点的空间距离就是两点之间一条线段的长度。

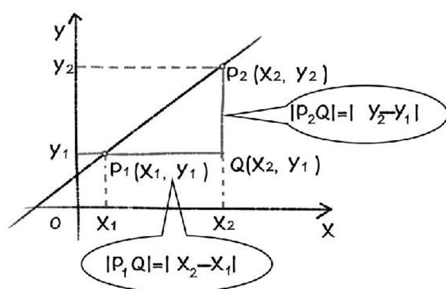
在一维空间中，两点$P_1$，（$x_1$）和$P_2$（$x_2$）之间的距离为：

$$s=|x_1-x_2|=\sqrt{(x_1-x_2)^2}$$



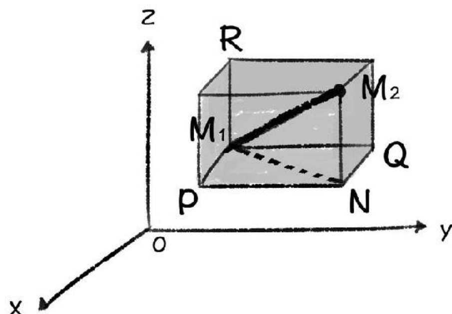

在二维空间中，两点$P_1$（$x_1$，$y_1$）和$P_2$（$x_2$，$y_2$）之间的距离可以用勾股定理计算，为：

$$s=\sqrt{(x_1-x_2)^2+(y_1-y_2)^2}$$




在三维空间中，两点$M_1$，（$x_1$，$y_1$，$z_1$）和$M_2$（$x_2$，$y_2$，$z_2$）之间的距离也可以通过两次勾股定理得到：

$$s=\sqrt{(x_1-x_2)^2+(y_1-y_2)^2+(z_1-z_2)^2}$$

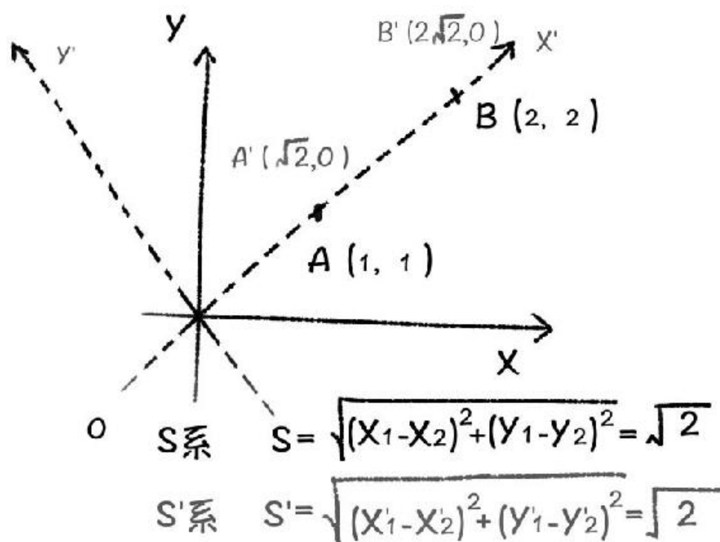



以此类推，在四维空间中，需要四个变量（$x$，$y$，$z$，$t$）才能表示一个点的坐标。两点（$x_1$，$y_1$，$z_1$，$t_1$）（$x_2$，$y_2$，$z_2$，$t_2$）之间的距离计算公式为：

$$s=\sqrt{(x_1-x_2)^2+(y_1-y_2)^2+(z_1-z_2)^2+(t_1-t_2)^2}$$

按照这种方法，我们甚至可以计算$N$维空间的两点间距离，这个距离称为“$N$维空间的空间间隔”。

空间间隔$s$是个很重要的量，如果坐标系选取不同，每个点的坐标就不同，但是空间距离保持不变。比如在平面直角坐标系中，我们采用实线坐标系和虚线坐标系，$A$和$B$两点的坐标都是不同的。但是，在这两个参考系下，计算的空间间隔都相等。这是因为线段$AB$的长度并不随着坐标系的变化而变化。



# 四、什么是降维打击

高维空间和低维空间有什么联系呢？其实，$N$ 维空间沿某维度的投影是$N-1$维空间。投影这个概念如果不好理解，我们可以理解为“切一刀”。比如：

如果我们在直线上切一刀，断面是一个点。

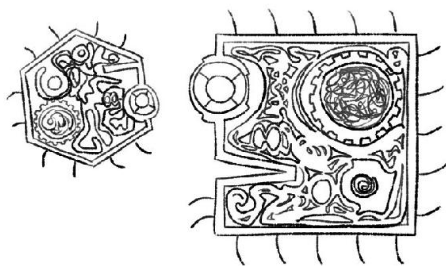
如果我们在平面上切一刀，断面是一条线。

如果我们在立方体上切一刀，断面是一个平面。

同样，我们在四维空间随便切一刀，都会切出三维空间。

刘慈欣的小说《三体》中谈到了降维打击，大体也是这个意思。把我们的三维空间沿着某条棱压缩，就会变为一个平面。就好比人家把我们压缩成平面人了，那我们不就死定了。

除了《三体》，还有许多其他科幻小说也对高维空间有讨论，如埃德温·阿伯特在他的书《平面国》中讲述了在一个扁平得就像一张纸的二维世界中的故事。在这个正方形的眼中，生活在三维世界中的人们拥有近乎神的力量，因为他们能在不打破（二维的）保险箱的情况下从其中把东西取出，能看到所有在二维世界看来是被挡在墙后面的东西，甚至能站在离二维世界几英寸的地方来保持“隐形”。



总而言之，高维空间与低维空间在数学上都没有什么区别。只是，我们使用三维空间就很好地解释了我们生活的世界，所以人们才感觉四维和四维以上的空间如此难以理解。我们必须明白，数学只是我们理解世界的工具，它本身并不需要为我们理解世界而负责。

# 1.12　数学家能在赌场中赢钱吗？

——概率论

我们经常听说这样的故事：一个数学家进了赌场，通过自己的专业能力赢了一大笔钱，那么现实生活中这样的事情可能吗？

十七世纪荷兰数学家、物理学家和天文学家惠更斯最早研究了赌场中的数学问题。他的著作《论赌博中的计算》使用概率论对赌博的结果进行分析，被认为是概率论的开端。




## 一、百家乐



在赌场中，有些赌博游戏有技术含量，但是更多的是全凭概率和运气的游戏。比如常见的赌场游戏：百家乐。这是一种发扑克牌比较大小的游戏，荷官会将八副牌全都混合在一起，然后给庄家和闲家各发2～3张牌。下注后开牌，庄家和闲家各自手中的牌的点数相加，尾数更大的

赢。

显然，这个游戏有三种可能：庄大、闲大，以及平局。玩家如果下注庄大或下注闲大，获胜之后，都会获得下注彩金等额的奖金（一赔一）。若下注平，会获得下注彩金9倍的奖金（一赔九）。那么，哪一种下注方法更容易赚钱呢？

这个问题比较复杂，我们把它简化为一个更加简单的骰子游戏。




游戏规则如下：

1．两个不透明的罐子分别属于庄家和闲家，里面各有一个骰子，每个骰子有1～6点。

2．摇晃后玩家下注，押庄大、闲大、平。

3．打开罐子比较，如果押大小获胜，一赔一，如果押平获胜，一赔五。

请问这个规则是否公平？

# 二、古典概型

解决这个问题，需要采用古典概型。所谓古典概型，就是一个过程共有$N$ 种可能的结果，而且每种结果发生的可能性相同。其中，事件$A$包含$M$ 种可能，那么事件$A$ 的概率就等于$M$ 和$N$ 的比，即：$P(A)=\frac{M}{N}$。

比如，一个班级里有50名同学，20名男生，30名女生，如果我随机抽取一个人，那么就有50种可能的结果。“抽到男生”这个事件包含20种可能的结果，因此抽到男生的概率就是$P$（男生）$=\frac{20}{50}=\frac{2}{5}$。

在这个游戏中，两个骰子都有6种可能的结果，因此两个骰子的组合有$N$ = 36种可能的结果，每种结果出现的概率都是1/36。

| | | 闲家点数 | | | | | |
|---|---|---|---|---|---|---|---|
| | | 1 | 2 | 3 | 4 | 5 | 6 |
| 庄家点数 | 1 | 平 | 闲大 | 闲大 | 闲大 | 闲大 | 闲大 |
| | 2 | 庄大 | 平 | 闲大 | 闲大 | 闲大 | 闲大 |
| | 3 | 庄大 | 庄大 | 平 | 闲大 | 闲大 | 闲大 |
| | 4 | 庄大 | 庄大 | 庄大 | 平 | 闲大 | 闲大 |
| | 5 | 庄大 | 庄大 | 庄大 | 庄大 | 平 | 闲大 |
| | 6 | 庄大 | 庄大 | 庄大 | 庄大 | 庄大 | 平 |

如果两个骰子的点数相同，那么就有1～1、2～2、…、6～6共计$M$ = 6种可能，因此出现平局的概率为：$P(\text{平})=\frac{M}{N}=\frac{6}{36}=\frac{1}{6}$。

也就是说，如果押平，赢的概率是1/6，输的概率是5/6。

# 三、数学期望

假设以1元钱下注，按照赔率，赢了就拿走6元钱，输了就不拿钱，平均能获得几元钱呢？

我们用输赢两种情况下的收益乘以相应的概率再相加，就可以计算出一个平均值。这个值在数学上称为“期望”。这里期望的意思就是，从统计意义上讲，每一次游戏结束后，平均可以拿回来的钱。

$$E=6\times\frac{1}{6}+0\times\frac{5}{6}=1$$

这说明：平均来讲，每局游戏以1元钱下注，平均可以拿走1元钱。即如果一直押平，而庄家不出千，那么概率上讲玩家是不赔不赚的。

我们再来看押大小。从刚才的表格中我们可以看出，庄大和闲大都包含15种可能，因此按照古典概型，任何一边大的概率都是：

$$P=\frac{15}{36}=\frac{5}{12}$$

也就是说，押两边任何一边大，赢的概率都是5/12，输的概率是7/12，还是假设以1元钱下注，那么押两边任何一边大，赢了都拿走2元，输了不拿钱，最后获得的数学期望都是：

$$E=2\times\frac{5}{12}+0\times\frac{7}{12}=\frac{5}{6}$$

我们会发现：押1元钱给任何一边大，统计意义上都会拿回5/6元，也就是说，平均会赔1/6元，是不合算的。

# 四、能利用数学赚钱吗？

也许有同学据此分析，只要一直押平，就可以不赔钱了。但是以上分析完全是建立在赌场不出千，纯凭概率的基础上。如果赌场出千，那么无论采取什么策略，玩家都会必输无疑。




还有的同学会说，既然我知道了各种情况下的概率，那么我可以通过观察来提高自己的胜率。例如，平局的概率是1/6，庄大的概率是5/12，如果连续2次开出平局，又连续开出5次庄大，那么接下来的5次就一定是闲大了。

这种说法是错误的，因为概率的意义是在开牌之前计算各种情况的可能性大小，一旦开牌，可能性就会变成确定性。而连续2次开牌之间并无实质联系，就算连续开了一百次平局，下一局是平局的概率依然是

1/6，而不会减小。

简单来说，数学会告诉你钱是怎么输的，但是不能帮助你赢钱。在电影《雨人》中，主角的哥哥患有自闭症，但是却具有超强的记忆力，靠着记忆力记下八副牌的顺序，赢了一大笔钱。而现实生活中是不可能的，因为荷官洗牌时并不会给你时间记牌，而当发牌到少于一定数目时，又会重新开始洗牌。想凭借数学或者记忆力在赌场里赚钱，是异想天开的。

# 1.13　买彩票能中大奖吗？

——排列、组合和概率

许多人在茶余饭后喜欢买张彩票，2元一注，中了奖当然好，中不了哈哈一乐就算了。也有人把买彩票当成了一份事业，花许多时间研究彩票，又花许多钱买彩票。彩票的走势真的有规律可循吗？长期买彩票到底是亏是赚？



## 一、排列三

我们首先来研究一个简单的彩票玩法：排列三，也叫作3D球。

这种彩票的基本玩法是：2元一注，从000～999选择一个数（三位），开奖时开出一个号码（三位），若下注的三位数和开奖的三位数数字和顺序都相同，即可中奖。否则不中奖。例如下注时选择052，开奖时第一个球是0，第二个球是5，第三个球是2，即可中奖。如果中奖，奖金1040元。

我们不妨从数学上计算一下中奖概率，再计算一下期望。

已知，开奖时每个球都有10种可能，三个球开出三个号码，排列起来，一共有1000种可能，也就对应着下注时000～999这1000个数字。

由于每个球开出的数字都是随机的，每种可能性概率都相等，所以这依然是一个古典概型。

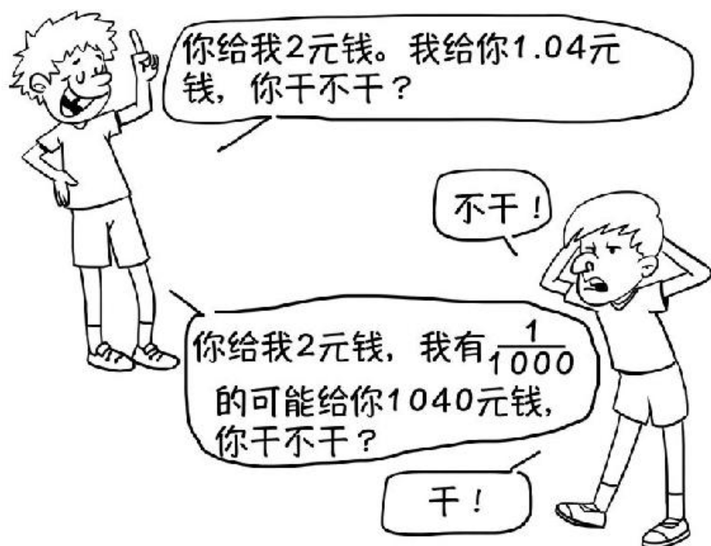
中奖包含多少种可能呢？因为只有一个中奖号码，因此中奖只有1种可能。

这样一来，中奖的概率就是$P=\frac{1}{1000}$。

中奖得1040元，概率1/1000；不中奖得0元，概率999/1000，所以玩一次返还的彩金期望是：

$$E=1040\times\frac{1}{1000}+0\times\frac{999}{1000}=1.04$$

也就是说，花2元下注，平均可以拿回1.04元，亏0.96元。

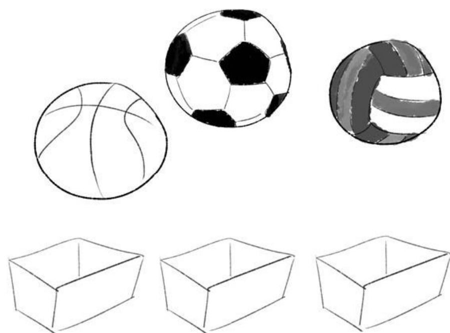



# 二、组选六

排列三还有另一种玩法，称为组选六，意思是说：开奖时的三个数字与下注的三个数字相同即可中奖，而不限次序。例如开奖是052，下注时选择052、025、520、502、250、205皆可中奖。中奖奖金173元。那么这种玩法的返奖期望又是多大呢？

已知，开奖结果依然是1000种可能，并且每种可能性概率都相同，古典概型。

中奖时，只要三个数字相同，而顺序可以随意调换。这就涉及一个概念：排列数。将$N$ 个物体按照一定的顺序进行排列，请问共有多少种排列方法呢？



比如，我们要把篮球、足球、排球放在有顺序的三个盒子里。首先研究第一个盒子，可以放任意1个球，所以有三种方法。在第一个盒子放好了球之后，第二个盒子就只有2个球可以选择，所以有两种方法。前两个盒子放好之后，只剩下一个盒子和1个球，因此最后一个盒子只有一种放球的方法。所以3个球放进三个盒子的方法数为$3\times2\times1=6$种。我们称3的全排列等于6。

同样，如果有$N$ 个球，放进有顺序的$N$ 个盒子里，称为“$N$ 的全排列”，公式为：

$$A_N^N=N\times(N-1)\times(N-2)\times\cdots\times1$$

我们下注一个三位号码之后，可以随意调换顺序，所以需要把这三个数字全排列，所以无论这三个数字是什么，都有6种排列方法。这样，中奖的可能就从一种变为了6种。

这样一来，中奖的概率就变成了$P=\frac{6}{1000}=\frac{3}{500}$。

中奖得173元，概率3/500；不中奖得0元，概率497/500，返奖期望为：

$$E = 173 \times \frac{3}{500} + 0 \times \frac{497}{500} = 1.038$$

也就是说，平均下注一次组选六，可以拿回1.038元，相比于直选玩法还少了0.002元。




# 三、双色球

相比于排列三，还有一种更让人热血沸腾的彩票玩法：双色球。

双色球有1～33共33个红球，还有1～16共16个蓝球。2元下注，在红球中选6个，蓝球中选1个。开奖时开出6个红球和1个蓝球，如果开出的红球与下注完全相同（不计顺序），蓝球也与下注相同（不计顺序），即可中大奖500万！




数学会告诉我们什么呢？为了研究这个问题，我们首先需要了解一个概念：组合数。

从$N$个不同的数字中选择$M$个，不计次序，一共有多少种选择呢？

比如，我们要从1、2、3、4、5、6六个数字中选三个，但是不计次序。我们可以按照下面的方法：

首先选一个数字，有6种选择方法。

在余下的数字中再选一个，有5种方法。

在余下的数字中再选一个，有4种方法。

所以一共的方法个数是：$6\times5\times4=120$种选择。

但是，如果选择的数字是$a$、$b$、$c$ 三个，按照刚才的算法，在120种可能中，$abc$、$acb$、$bca$、$bac$、$cab$、$cba$ 算作了不同的情况。按照不计次序的规则，这6种情况应该算作一种。于是，我们必须将120除以6，才能得到不计次序的选择方式个数：20种。这里除掉的6其实就是3的全排列。

那么同理，从$N$ 个数字中选择$M$ 个数字，首先从$N$ 开始乘，乘$N$ -1，乘$N$ -2，…，一共乘$M$ 个数字，得到计次序时的结果。然后再除以$M$ 的全排列，得到不计次序的结果。公式写作：

$N$中取$M$计算次序的可能

$$C_N^M=\frac{N(N-1)(N-2)\cdots(N-M+1)}{1\times2\times3\times\cdots\times M}$$

$N$中取$M$不计次序的可能

$M$个数字排列的可能

明白了这个算法，我们就可以计算双色球中大奖的概率了。

首先计算开奖结果一共有多少种可能。从33个红球中选6个，可能的结果有：

$$C_{33}^6=\frac{33\times32\times31\times30\times29\times28}{1\times2\times3\times4\times5\times6}=1107568$$

从16个蓝色球中选1个，显然有16种可能结果。

所以，开奖结果一共的可能数为$1107568\times16=17721088$。

也就是说，双色球的开奖结果大约有1700万种可能。而且，每一种开奖结果的可能性都是相同的。如果我们花2元钱买一注，只能买一种可能，因此中奖的概率为大约1700万分之一。

我们来计算一下数学期望。中大奖能拿到500万，扣掉20%的税，还剩下400万，于是数学期望为：

$$E=400万\times\frac{1}{1700万}+0\times\frac{1700万-1}{1700万}=0.235$$

也就是说，如果只奔着大奖去，平均买一注彩票，只能拿回0.235元。

当然，双色球还有二三四五六等奖，每一种奖的奖金和对应概率都不同，尤其是一等奖和二等奖，采用浮动奖金制度，计算起来更加烦琐。根据中国福利彩票发行管理中心公布的中奖规则，双色球返奖率大约为50%，也就是说，花2元钱下注，大奖小奖全算上，统计上讲，平均可以拿回1元钱，而另1元钱就是给福利事业做贡献了。

# 1.14　天气预报为啥总不准？

——条件概率

许多人说，现在科学这么发达，为什么天气预报总不准呢？这里涉及一个数学问题，称为“条件概率”。




什么是条件概率呢？比如我们要确定6月15日是不是下雨，根据往年数据，下雨的概率有40%，不下雨的概率为60%，这就称为“概率”。如果在前一天，天气预报说6月15日下雨，这就称为“条件”，在这种条件下，6月15日真正下雨的概率就称为“条件概率”。

## 一、下雨和不下雨

天气预报根据一定的气象参数推测是否会下雨，由于天气捉摸不定，即便预报下雨，也有可能是晴天。假设天气预报的准确率为90%，即在预报下雨的情况下，有90%的概率下雨，有10%的概率不下雨；同样，在预报不下雨的情况下，有10%的概率下雨，有90%的概率不下雨。

这样一来，6月15日的预报和天气就有四种可能：预报下雨且真的下雨，预报不下雨但是下雨，预报下雨但是不下雨，预报不下雨且真的

不下雨。我们把四种情况列在下面的表格中，并计算相应的概率。

| | 下雨 | 不下雨 |
|---|---|---|
| 预报下雨 | 40% × 90% = 36% | 60% × 10% = 6% |
| 预报不下雨 | 40% × 10% = 4% | 60% × 90% = 54% |

计算方法就是两个概率的乘积。例如下雨概率为40%，下雨时预报下雨的概率为90%，因此预报下雨且下雨这种情况出现的概率为36%。同样，我们可以计算出天气预报下雨但是不下雨的概率为6%，二者之和为42%，这就是天气预报下雨的概率。

在这42%的可能性中，真正下雨占36%的可能，比例为$\frac{36\%}{42\%}\approx 85.7\%$，而不下雨的概率为6%，占$\frac{6\%}{42\%}\approx 14.3\%$。也就是说，假设天气预报的准确率为90%，预报下雨的条件下，真正下雨的概率只有85.7%。

我们会发现：预报下雨时是否真的下雨，不光与预报的准确度有关，同时也与这个地区平时下雨的概率有关。




# 二、生病和没生病

与这个问题类似的是在医院进行重大疾病检查时，如果医生发现异

常，一般不会直接断定生病了，而是建议他去大医院再检查一次，虽然这两次检查可能完全是一样的。为什么会这样呢？

我们假设有一种重大疾病，患病人群占总人群的比例为1/7000。也就是说，随机选取一个人，有1/7000的概率患有这种疾病，有6999/7000的概率没有患这种疾病。

有一种先进的检测方法，误诊率只有万分之一，也就是说，患病的人有1/10000的可能性被误诊为健康人，健康人也有1/10000的可能性被误诊为患病。

我们要问：在一次检查得到患病结果的前提下，这个人真正患病的概率有多大？

| | 患病 | 健康 |
|---|---|---|
| 检测患病 | $\frac{1}{7000}\times\frac{9999}{10000}=\frac{9999}{70000000}$ | $\frac{6999}{7000}\times\frac{1}{10000}=\frac{6999}{70000000}$ |
| 检测健康 | $\frac{1}{7000}\times\frac{1}{10000}=\frac{1}{70000000}$ | $\frac{6999}{7000}\times\frac{9999}{10000}=\frac{69983001}{70000000}$ |

我们仿照刚才的计算方法，检测出患病的总概率为：

$$\frac{9999}{70000000}+\frac{6999}{70000000}=\frac{16998}{70000000}$$

而患病且检测出患病的概率为$\frac{9999}{70000000}$。

所以在检测患病的条件下，真正患病的概率为$\frac{9999/70000000}{16998/70000000}=\frac{9999}{16998}\approx58.8\%$。

显而易见，即便是万分之一误诊的情况，一次检测也不能完全确定这个人是否患病。




那么，两次检测都是患病的情况又如何呢？

大家要注意，在第一次检测结果为患病的前提下，此人患病的概率已经不再是所有人群的1/7000，而变为了自己的58.8%，健康的概率只有41.2%，此处的概率就是条件概率。所以第二次检测的表格变为：

| | 患病 | 健康 |
|---|---|---|
| 检测患病 | $58.8\% \times \frac{9999}{10000} \approx 58.794\%$ | $41.2\% \times \frac{1}{10000} \approx 0.004\%$ |
| 检测健康 | $58.8\% \times \frac{1}{10000} \approx 0.006\%$ | $41.2\% \times \frac{9999}{10000} \approx 41.196\%$ |

在两次检测都是患病的条件下，此人真正患病的概率为：

$$\frac{58.794\%}{58.794\%+0.004\%} \approx 99.99\%$$

基本确诊了。

# 三、贝叶斯公式

对这个问题进行详细讨论的人是英国数学家贝叶斯。



贝叶斯

贝叶斯指出：如果$A$ 和$B$ 是两个相关的事件。$A$ 有发生和不发生两种可能，$B$ 有$B_1$、$B_2$、…、$B_n$ 共$n$ 种可能。那么在$A$ 发生的前提下，$B_i$ 发生的概率称为条件概率$P(B_i|A)$。

要计算这个概率，首先要计算在$B_i$ 发生的条件下$A$ 发生的概率，公式为$P(B_i)P(A|B_i)$。

然后，需要计算事件$A$ 发生的总概率，方法是用每种$B_i$ 情况发生的概率与相应情况下$A$ 发生的概率相乘，再将乘积相加。

$$P(B_1)P(A|B_1)+P(B_2)P(A|B_2)+\ldots+P(B_n)P(A|B_n)$$

最后，用上述两个概率相除。完整的贝叶斯公式就是：

$$P(B_i|A)=\frac{P(B_i)P(A|B_i)}{P(B_1)P(A|B_1)+P(B_2)P(A|B_2)+\cdots+P(B_n)P(A|B_n)}$$

贝叶斯公式在社会学、统计学、医学等领域，都发挥着巨大的作用。

下次遇到天气误报、医院误诊，不要完全怪气象台和医院啦，有时候，这是个数学问题。

# 1.15　散户炒股为啥总赔钱？

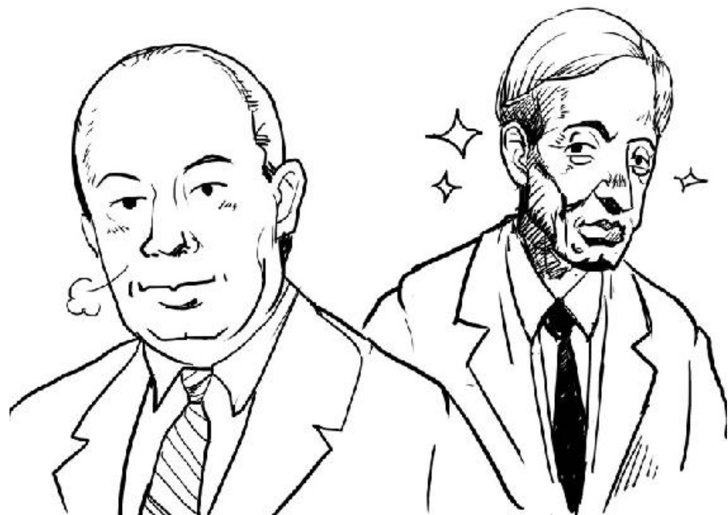
——博弈论基础

不知道各位同学中有没有炒股的，虽然股市不停地在牛熊市之间转换，但是散户大多数都是赔钱的。关于炒股赔钱这件事，有人认为是自己智商不够，有人认为是自己运气不好。那么，在股市里有没有什么更深刻的数学内涵呢？

股市类似于一种零和游戏，每个人都希望别人赔钱，而自己赚钱。这就涉及一个数学过程：博弈。博弈的本意是下棋，现在引申为通过一定的策略，使自己的利益最大化。




博弈论最早是由计算机之父冯·诺依曼提出的，后来经过约翰·纳什等人发扬光大。博弈论不同于概率论，博弈论是指参与者可以主动调整自己的策略，从而获得最大收益，它是现代数学的一个分支，在金融、政治、计算机等领域都有广泛应用。



# 一、美女与男人的硬币游戏

为了理解博弈论，我们来举一个经典的例子，称为美女与男人的游戏。

有一个男人在酒吧里喝酒，一位美女走过来，对他说：我们玩个游戏吧。规则如下：

1．每个人手里各拿一个硬币，扣在桌子上（不让对方看到）。

2．两人同时把手拿开，看硬币的正反面。

3．如果硬币都是正面，那么美女给男人3块钱。如果都是反面，那么美女给男人1块钱。如果硬币是一正一反，男人给美女2块钱。



如果从概率论的角度来考虑，很容易感觉这是一个公平的游戏。因为如果两个人都是随机扣下硬币，那么两个都正面的概率为1/4，两个都反面的概率为1/4，一正一反的概率为1/2，那么男人收益的情况如下：

| 男人收益 | | |
|---|---|---|
| | 3 | −2 |
| | −2 | 1 |

按照概率来计算，一次游戏中，男人收益的数学期望就是：

$$E=\frac{1}{4}\times 3+\frac{1}{4}\times 1+\frac{1}{2}\times(-2)=0$$

也就是说，经过一次游戏，男人的平均回报为0，既不赚钱，也不

亏钱。

事实真的如此吗？

# 二、美女会采用什么策略？

在这个游戏中，男人和女人并不是抛硬币，而是自己选择出硬币正面还是反面。显然，男人和女人都不可能一直出正面或者一直出反面，因为这样会被对手摸出规律。但是他们依然可以在多次游戏中将自己出正面的频率设定在某个值附近，从而获得统计意义上的收益，这就使得游戏从一个概率问题，变成了一个博弈问题。

为了解决这个问题，我们假设男人出正面的频率为$x$，则男人出反面的频率为$1-x$；设女人出正面的频率为$y$，则女人出反面的频率为$1-y$。各种情况出现的频率如表格所示。

| 概率频率 | | |
|---|---|---|
| | $xy$ | $(1-x)y$ |
| | $x(1-y)$ | $(1-x)(1-y)$ |

结合男人的收益表格和频率表格，男人的收益期望等于各种情况出现的概率与收益的乘积之和。

所以，男人在一次游戏后的数学期望是：

$$E = 3xy - 2(1-x)y - 2x(1-y) + (1-x)(1-y) = 8xy - 3x - 3y + 1$$

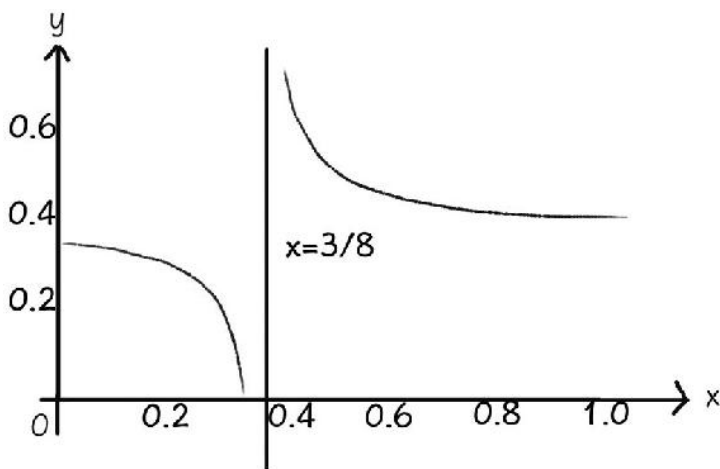
博弈双方对这个结果是有不同预期的。女人希望男人一直赔钱，所以希望男人的收益期望是负的。而男人希望自己一直赢钱，所以希望自己的收益期望是正的。两人可以采用的策略就是调整自己出正面的频率$x$，$y$，这两个频率都在0和1之间。

因为游戏是女人设计的，所以我们不妨从女人的角度考虑。女人希望调整自己的正面频率$y$，使得无论男人的正面频率$x$为多大，不等式$8xy-3x-3y+1<0$总是成立的。那么她能达到目的吗？

我们做一个移项，得到$(8x-3)y<3x-1$，这个不等式要分为三种情况：

1．若$8x-3=0$，即$x=\frac{3}{8}$，则不等式变为$0<\frac{1}{8}$，显然，这是永远成立的。

2．若$8x-3>0$，即$x>\frac{3}{8}$，则不等式变形为$y<\frac{3x-1}{8x-3}$。要保证这个不等式一直成立，就需要$y$小于$\frac{3x-1}{8x-3}$的最小值。




画出$\frac{3x-1}{8x-3}$的图像，我们会发现：在$x>\frac{3}{8}$和$x<\frac{3}{8}$两个范围内，函数都是单调递减的。

当$x>\frac{3}{8}$时，函数单调递减，所以$x=1$时，$\frac{3x-1}{8x-3}$取得最小值，最小值为$\frac{3-1}{8-3}=\frac{2}{5}$。所以在$x>\frac{3}{8}$时，使不等式一直成立的解为$y<\frac{2}{5}$。

3．若$8x-3<0$，则由于变号原因，不等式变形为$y>\frac{3x-1}{8x-3}$。要保证这个不等式一直成立，就需要$y$ 大于$\frac{3x-1}{8x-3}$的最大值。根据图像，在$x<\frac{3}{8}$时，$x$ 越小，函数值越大，$x=0$时，函数取得最大值，即$\frac{3x-1}{8x-3}=\frac{1}{3}$。也就是说，当$x<\frac{3}{8}$时，$y>\frac{1}{3}$才能保证不等号永远成立。




综上所述，当$y$ 的取值在$\frac{1}{3}\sim\frac{2}{5}$之间时，无论$x$ 是多大，不等式都是成立的。也就是说，如果女人的硬币出正面的频率在$\frac{1}{3}\sim\frac{2}{5}$之间，那么无论男人采取什么策略，他的收益期望都是负的，统计意义上一定会赔钱。

# 三、庄家是如何收割散户的？

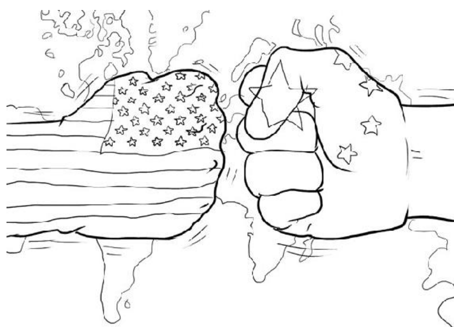
这是不是很像股市？在股市中，庄家可以操纵股价上下翻飞，让你的心痒痒的，就好像美女一般。在庄家拉升股价时，我们做多，就可以盈利。庄家打压股价时，我们做空，也可以盈利。但是如果庄家做多，我们做空，或者相反，就会亏损。在这样的规则下，每个人都觉得自己可能是个幸运儿，可以通过自己的运气或者策略获得正的收益。

但是事实上，庄家有比散户更强的控盘能力和模型计算能力，他们会采用一种更好的策略，使得散户无论采取什么方式炒股，统计意义上都会赔钱。当然，不排除有些散户的运气特别好，在一段时间内大赚了一笔。但即便如此，我们依然要说，在这样的规则下，长期炒股的散户，才会多数都赔钱。

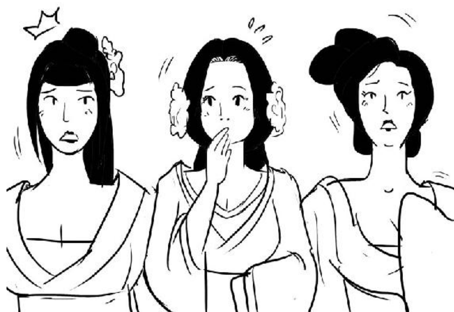
# 1.16　老板为啥对底层员工特别好？

——再谈博弈论

不知道大家发现没有，公司老板总是对底层员工态度特别好，就算员工犯了错，老板也是和风细雨。但是老板对自己的副手以及中层领导就没那么客气了，有时候还会随时提防着。这是为什么呢？



## 一、三姬分金



为了能够通俗地理解这个问题，我们不妨从一部动画片《天行九歌》中的桥段《三姬分金》说起。

在这个桥段中，韩非子去找大将军姬无夜筹措军饷。

发现大帐之中除了将军外，还有三名美女在玩抢金币的游戏。韩非子对三位美女说，咱们不妨玩得更有趣一些。新的游戏规则是：

1．抽签决定三个人的顺序A、B、C，按照顺序进行分金币的提议；

2．如果提议未能获得全体人员半数以上（不包括半数）通过，提议人被处死，由下一个人提议；

3．如果提议获得全体人员半数以上通过，按该提议分金币，游戏结束。

在这个游戏规则下，抽到第一名提议的美女非常恐慌，因为她觉得后面两个人为了拿更多的金币，必然会否定自己的提议，然后杀死自己。但是情况真的是这样吗？

# 二、博弈策略

为了使用博弈论分析这个问题，首先我们必须做出几点假设：

1．美女都是聪明的，知道自己的决策会导致什么结果；

2．美女都是理性的，以自己的利益最大化为目标；

3．美女都是邪恶的，在利益最大化的前提下，尽量多杀人。

在这样的假设下，我们开始讨论这个问题。

1．首先假设$A$ 已经被杀了，那么只剩下$B$ 和$C$ 两个人，此时无论$B$提出什么建议，$C$ 都可以反对，这样提议没有获得半数以上支持，$B$ 被杀死。$C$ 不光可以拿到全部金币，还杀掉了两个人，$C$ 获得利益最大；

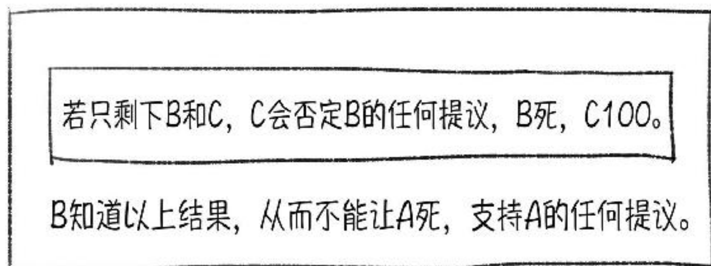
2．$B$ 知道以上结果，所以$B$ 的策略是绝对不能让$A$ 死掉，转而支持$A$ 的一切提议；

3．$A$ 知道以上结果，有$B$ 的支持，$A$ 自己也支持自己，所以$A$ 的任

何提议都会被通过，因此$A$ 的提议是$A$ 100，$B$ 0，$C$ 0。此时$C$ 反对已经没有任何意义了。

最终$A$ 拿到了全部的金币，$B$ 和$C$ 什么都拿不到。

我们可以使用框图来表示这个过程。




# 三、四个人玩，结果如何？

我们不妨设想，如果四个人玩这个游戏，结果又是如何呢？如果大将军姬无夜$M$ 也要玩这个游戏，并且$M$ 第一个提议，他会知道以上结果。他知道如果自己死掉，那么$A$ 会分走全部的金币，而$B$ 和$C$ 什么都拿不到。而且，四个人要有超过半数同意自己，至少需要三个人支持，除了自己之外，他还需要拉拢两个人。

显然，拉拢$B$ 和$C$ 更好。因为如果自己死掉，$B$ 和$C$ 什么都拿不到，于是只要$M$ 给$B$ 和$C$ 每人一个金币，自己拿98个，$B$ 和$C$ 就一定支持自己，此时$A$ 反对已经没有任何意义了。

所以$M$ 的提议会是$M$ 98，$A$ 0，$B$ 1，$C$ 1。用框图表示如下：

若只剩下B和C,C会否定B的任何提议，B死，C100。

B知道以上结果，从而不能让A死，支持A的任何提议。

A知道以上结果，从而提议A100，B0，C0。

M知道以上结果，要拉拢B和C，从而提议M98，A0，B1，C1。

有人可能会想，$A$ ，$B$ ，$C$ 为什么不联合起来把$M$ 干掉，约定干掉之后，她们每人拿33个金币？的确，她们可以这样做，但是当$M$ 被干掉之后，就面临一个问题：$A$ 会不会反悔呢？假如$M$ 死了，$A$ 反悔了，提议自己拿100个，$B$ 和$C$ 还是什么也拿不到。

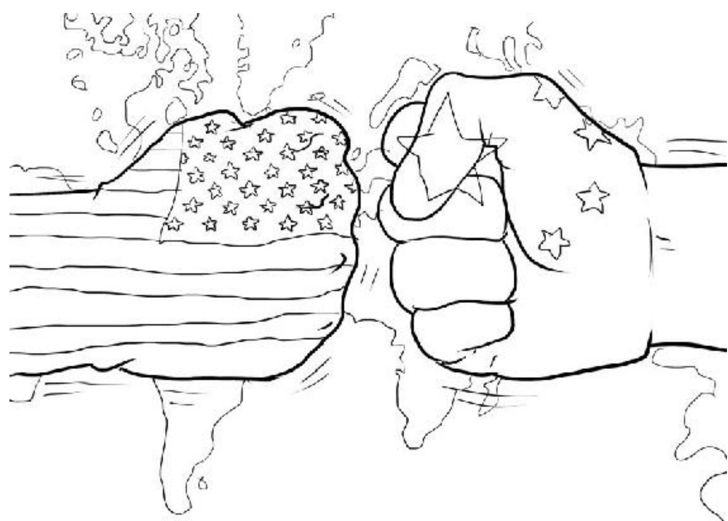
当然，$B$ 和$C$ 此时也可以联合起来把反悔的$A$ 干掉，然后约定每人拿50个金币。但是如果$A$ 死了，$C$ 又会不会反悔呢？如果$C$ 反悔了，$B$ 一定会死。

因为每个人都是理性的，又是邪恶的，他们不会相信其他人的承诺，不敢冒这个风险，所以$M$ 的分配关系还是会通过。

# 四、现实生活中的博弈问题

在现实生活中，这样的例子比比皆是。$M$ 就像是一个大公司的老板，他具有先手优势，因此可以为自己谋取最大的利益。$B$ 和$C$ 属于底层员工，他们比较安全，但是收益很少。不过，$M$ 特别喜欢拉拢$B$ 和$C$ ，就好像很多公司老板都对底层员工特别照顾，总是施以小恩小惠一样，因为他们是最好拉拢的。

但是$A$ 的位置很尴尬，他既没有先手优势，也不属于大老板拉拢的对象。他要获得最大利益，就必须干掉$M$ ，自己成为先手。所以历史上臣弑君、君杀臣的现象屡见不鲜。



国家之间的关系也是一样。例如美国作为世界老大，总是联合一些三四流国家整老二。当年的苏联是老二，美国通过意识形态把苏联搞散了；后来日本是老二，广场协定把日本搞残了；再后来俄罗斯越来越厉害，通过石油，把俄罗斯搞废了；现在中国是老二，美国又通过贸易战开始搞中国。因为他担心老二总想取代自己的位置。

博弈论是一种数学结论，在一定的假设条件下成立。现实生活远远比模型要复杂，所以，请不要把数学结论死套在生活中，也不要用生活中的个别案例来否定数学。

# 1.17　为啥总有人开车加塞？

——纳什均衡

北京的道路情况非常糟糕，下点小雨或者有个事故，就会堵成一锅粥。主要原因是车太多，但是也有一部分原因是有些司机喜欢加塞。本来可以走的路，就都堵死了。

为什么许多人喜欢开车加塞呢？这其实也是一个数学问题：纳什均衡。




纳什均衡是博弈论中的一种情况，指的是在一个博弈过程中，博弈双方都没有改变自己策略的动力，因为单方面改变自己的策略都会造成自己收益减少。纳什均衡点可以理解为个体最优解，但并不一定是集体最优解。

为了解释这个问题，我们举两个最典型的例子：囚徒困境和智猪博弈。

## 一、囚徒困境

囚徒困境是说：有两个小偷集体作案，然后被警察捉住。



警察对两个人分别审讯，并且告诉他们政策：

如果两个人都坦白交代作案过程和赃物去向，就可以定罪，两个人各判8年。

如果一个人交代，另一个不交代，那么一样可以定罪。但是交代的人从宽处罚，批评教育就释放。不交代的人从严处罚，判10年。

如果两个人都不交代，没法定罪，每个人判1年意思一下。

我们把两个人的收益情况写在表格里。由于判刑是一种惩罚，所以收益写作负的。

| | B坦白 | B抗拒 |
|---|---|---|
| A坦白 | −8，−8 | 0，−10 |
| A抗拒 | −10，0 | −1，−1 |

首先我们考虑$A$ 的决策。$A$ 会想，我如何才能获得更大收益呢？如果$B$ 坦白了，那么我坦白就会判8年，抗拒就会判10年，为了让自己的收益更大，我应该坦白；如果$B$ 抗拒了，我坦白会判0年，我抗拒会判1年，我还是应该坦白。所以最终$A$ 会选择坦白。

同样，$B$ 也会这样想，因此最终两个人都会坦白，每个人都判8年。

而且在两人都坦白的情况下，没有任何一方愿意单方面改变决策，因为一旦单方面改变决策，就会造成自己的收益下降。这个都坦白的点就称为“纳什均衡点”。

纳什均衡

| | B坦白 | B抗拒 |
|---|---|---|
| A坦白 | −8，−8 | 0，−10 |
| A抗拒 | −10，0 | −1，−1 |

B改变决策

A改变决策

显然，集体最优解是两个人都抗拒，这样一来，每个人都判1年就出来了。但是，纳什均衡点却不在这里。这就说明个人理性的结果未必是集体最优解。

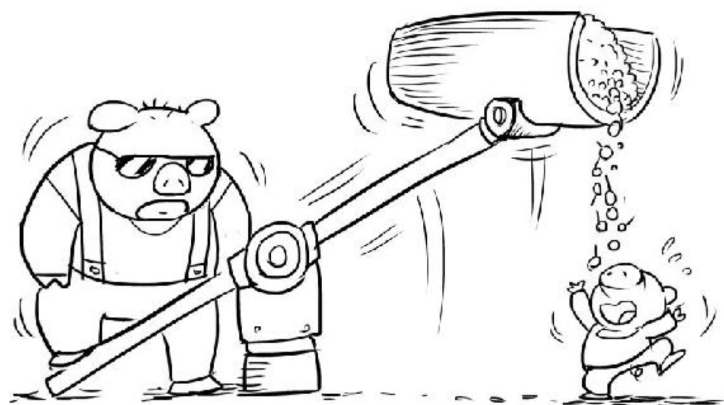
这与我国开车加塞的例子很像。如果大家都不加塞，是整体的最优解，但是按照纳什均衡理论，任何一个司机都会考虑：无论别人是否加塞，我加塞都可以使自己的收益变大。于是最终大家都会加塞，加剧拥堵，反而不如大家都不加塞走得快。

那么，有没有办法使个人最优变成集体最优呢？方法就是共谋。两个小偷在作案之前可以先说好，咱们如果进去了，一定要选择抗拒。如果你这一次敢反悔，那么以后道上就再也不会有人跟你一起了。也就是说，在多次博弈过程中，共谋是可能的。但是如果这个小偷想干完这一票就走，共谋就是不牢靠的。

在社会领域，共谋是靠法律完成的。大家约定的共谋结论就是法律，如果有人不按照约定做，就会受到法律的惩罚。通过这种方式保证最终决策从个人最优的纳什均衡点变为集体最优点。

# 二、智猪博弈

智猪博弈是另一个纳什均衡的典型例子：一个食槽中装有10份食物，但是控制按钮在另一端，需要到另一端按下按钮，食物才能掉下来。大猪和小猪都在食槽一端，它们两个都可以跑到另一端按下按钮再回来，二者速度相同，消耗相同的体力，并且一只猪跑去按按钮，会造成另一只猪先吃食物。



我们假设每只猪跑去按按钮都要消耗2份食物的体力，并且大猪比小猪吃食物快：

如果大猪先吃食物，二者吃食物的比例为9：1；

如果小猪先吃食物，二者吃食物的比例为6：4；

如果二者同时吃食物，吃食物的比例为7：3。

两只猪都可以选择去按按钮，也可以选择等待。比如大猪去按按钮，小猪等待，那么小猪先吃食物，二者吃的食物比例为6：4。但是大猪消耗2份体力，所以最终大猪收益为4，小猪不消耗体力，收益为4。按照这种方法，我们可以写出各种决策时两只猪对应的收益：

| | 小猪去 | 小猪等 |
|---|---|---|
| 大猪去 | 5，1 | 4，4 |
| 大猪等 | 9，−1 | 0，0 |

我们来考虑均衡点。

小猪会思考：如果大猪去，我跟着去获得收益1，我等着获得收益4，因此我应该等着；如果大猪不去，而我去，我获得收益-1，如果我们都等着，我收益为0，因此我还是应该等着，这样一来，小猪的决策一定是等待。

在小猪等待的情况下，如果大猪去按按钮，获得收益4，如果大猪不去按按钮，获得收益0，因此大猪会选择去按按钮。这个（4，4）的收益就是纳什均衡点。

这和国家或者公司进行基础研究研发新产品很像。比如一款新的芯片研发需要花很多钱，成功后也能获得更大的收益。在这样的情况下，小国家小公司是没有动力进行研发的，他们会等待大国家大公司研发好了之后，直接利用现成的技术获得收益。

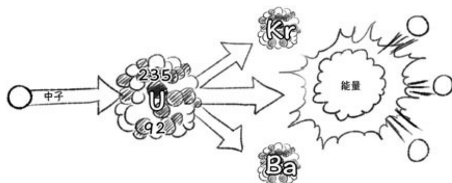
我们的芯片产业就是这样一个局面，多年以来，我们一直认为自己是发展中国家，没有大力推动半导体产业的基础研究，许多人秉着“做不如买，买不如租”的观点。现在美国对中国展开贸易战，禁止芯片出口，一下子就击中了一处弱点。

发现“纳什均衡”的约翰纳什是一位传奇人物，他是位数学家，却获得了诺贝尔经济学奖。晚年时精神分裂。想了解纳什的一生，可以去看看电影《美丽心灵》，这部电影拍得不错。

# 第二章　奇妙的物理

P—H—Y—S—I—C—S

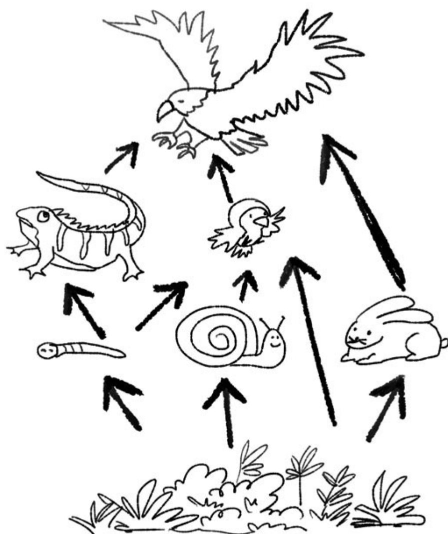
P



# 2.1　能量都是从哪儿来的？

——能量的转化与守恒

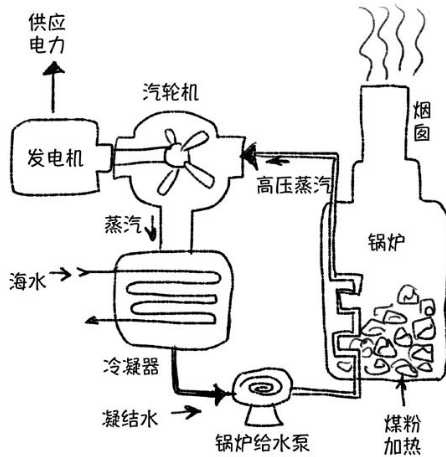
我们都知道一个概念：能量。用日光灯需要电能；开汽车需要石油的化学能；动物能够活动，因为具有生物能。那么能量到底是从哪里来的呢？



我们不妨从常见的电能来解释一下能量的来源。我们的家用电是由发电厂发出来的。发电厂为了发电，必须使发电机转动起来。要使发电机转动起来，就必须消耗其他能量。

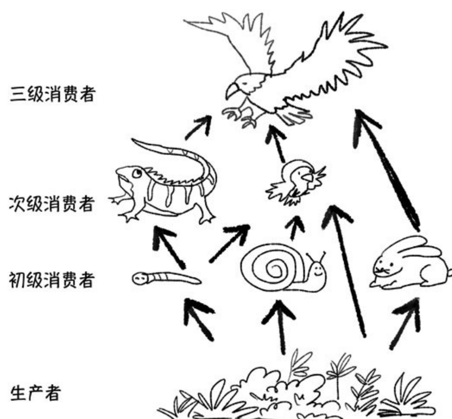
## 一、煤和石油的能量

我们首先看火力发电。




火力发电机的基本原理是：燃烧煤粉将水变为高压蒸汽，通过高压蒸汽推动汽轮机，汽轮机转动带动发电机转动，这样就可以产生电能。在这个过程中，煤的化学能被燃烧消耗，转化为机械能，机械能转化为电能。

煤的化学能是从哪里来的呢？煤和石油是千百年前动植物尸体在一定条件下形成的。详细地说，动植物死亡后，尸体沉积在地下，如果微生物分解尸体的速度没有沉积速度快，那么动植物尸体就会越来越多。在一定条件下，这些尸体中的有机物会发生变化，经过相当长的时间变成煤和石油。所以，煤和石油的化学能是从遥远古代的生物能转化而来的。



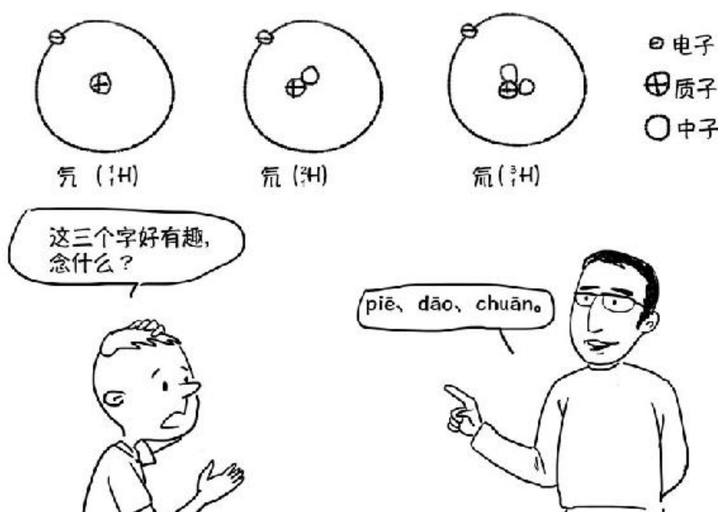

那么，生物的能量又是从哪儿来的呢？我们知道，食肉动物吃食草动物，食草动物吃植物，这就是食物链。食物链上级消耗者的生物能

（食肉动物）来源于下级能量消耗者（食草动物）和生产者（植物）。也就是说，动植物的生物能本质上来源于植物。

那么，植物的能量是从哪儿来的呢？除了蘑菇等依靠腐殖质生存的植物之外，大部分植物都有叶绿素。叶绿素可以进行光合作用。所谓光合作用，就是通过太阳能，将空气中的二氧化碳和吸收的水分转化成有机物，同时释放氧气的过程。在这个过程中，植物的生物能变多，看似产生了能量，所以称为生产者。但是实际上，植物必须利用太阳能才能完成这个过程。所以依然是能量的转化：把太阳能转化为了有机物的能量，或者称为“生物能”。

我们继续想，太阳能又是从哪里来的呢？

太阳是一个大火球，每时每刻都在释放着巨大的能量。太阳之所以能发光发热，是因为太阳内部有大量的氢元素。氢原子是自然界最小的原子，原子核中只有一个质子，原子核外有一个电子。但是，氢原子核里的中子个数可以不同，据此可以把氢元素分为三种同位素，分别是有没有中子的氕、一个中子的氘、两个中子的氚。

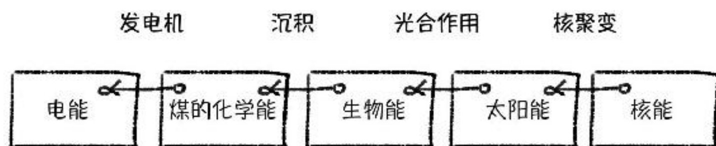



在极高的温度下，氘和氚原子核可以发生聚变反应，生成一个氦原子核和一个中子，同时释放巨大的能量。于是，这个能量就通过辐射的形式散发到宇宙中。

所以，太阳能来源于太阳内部的核聚变反应——核能。

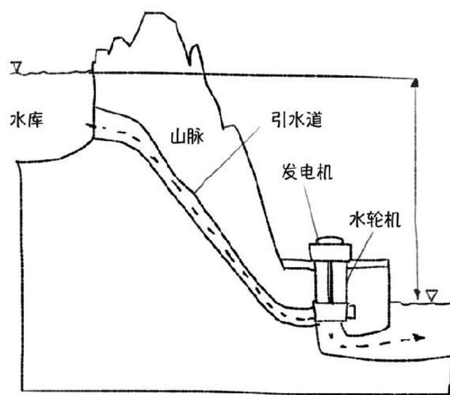
核能是从哪里来的？这个问题人类还没有理解。因为在大爆炸之初，核能就存在于宇宙之中了。而大爆炸之前的事，科学家们还没有任何头绪。

总结起来，火力发电的能量转化过程是太阳核能通过聚变转化为光能，光能通过光合作用转化为植物的生物能，植物的生物能通过食物链布满生物圈，生物能通过一定的环境形成煤的化学能，化学能通过燃烧加热蒸汽推动发电机，继而发电机就把这种能量转化为电能。




# 二、风和水的能量

如果我们用水电站发电，能量又是从哪儿来的呢？




要利用水能发电，首先要在有山脉的地区建设水坝，让水形成落差。然后，水受到重力作用向下运动，可以推动水轮机转动，这样发电机就可以工作了。由于水在高处可以向下流动带动发电机，于是我们就说高处的水具有能量，这种能量称为“重力势能”。水电站是把水的重力势能转化为电能。

水从高处到达低处之后，重力势能减小，高山上流下来的水汇集到一起，最终流到大海中去。那么，高山上的水为什么不会流干呢？这又涉及了自然界水的循环。

由于太阳的照射，海洋水可以蒸发成气体，升到空中。在风的带动下，重新来到高山上空形成降雨，这样水就可以从低处跑回到高处，重力势能增加。但是在这个过程中，太阳的照射必不可少，因为蒸发需要吸热，而这个热量来源就是太阳光。所以，水能也是太阳能转化而来的。




同样，风力发电是利用了风的能量，而风的能量同样是源于太阳能。太阳的能量中，只有22亿分之一到达了地球，可这22亿分之一却把地球变成了一个欣欣向荣的世界。

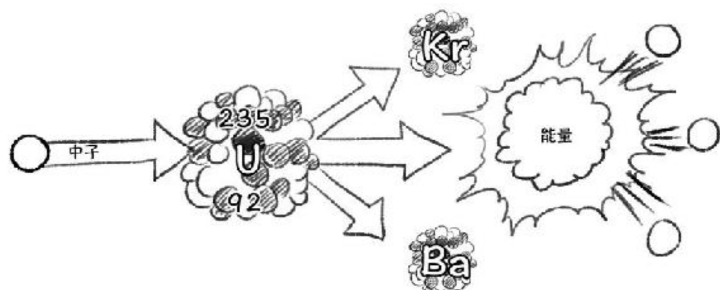
# 三、除了太阳还有其他能量来源吗？

地球上大部分能量都来源于太阳。但是，还有其他两种能量来源。

比如核电站。目前，像太阳那样的核聚变能量，人类只能用来制造氢弹，还没有掌握能够控制核聚变反应速度的技术用它来发电。但是人们已经掌握了可控核裂变技术。

核裂变是指重原子核可以裂变成较轻的原子核，同时释放能量的过

程。比如，铀235就是一种常见的核反应原料。当一个中子撞击铀235原子核时，会裂变出氪原子核Kr和钡原子核Ba，同时释放出三个中子。三个中子再撞击三个铀235，可以释放出9个中子……这个过程称为“链式反应”。链式反应可以释放巨大的能量。

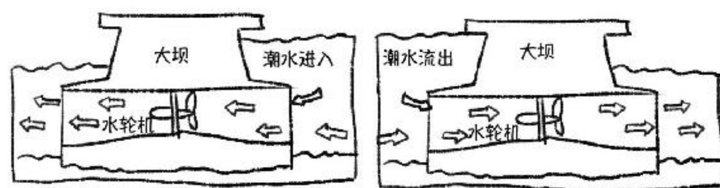



人们建设核反应堆，让链式反应缓慢地进行，放出大量的热把水变为高压蒸汽，推动发电机发电，这就是核电。

相比于火电，核电带来的污染小得多，所以很多国家都在大力发展核电。核电的能量来源是地球上储存的铀235，不是太阳。此外，还有一种能量不是来源于太阳——地热，地热来源于地球内部的岩浆，而岩浆的能量来源是地球内部的核裂变反应。

能量还有一种来源是月球。由于太阳和月球的引力作用，地球上的水会发生涨潮和落潮，而月亮对地球上潮汐的影响更大。

如果在潮水中安装发电机，潮水就可以推动发电机发电，这就是潮汐电站的原理。潮汐能从本质上讲是地球和月球具有的机械能转化而来的，这种机械能在地球和月球形成之初就存在了。



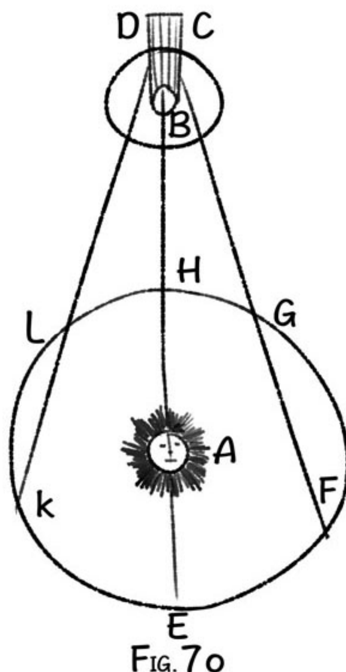

说了这么多，大家是不是已经明白了，能量总是在相互转化。而地球上绝大多数的能量来源有三个：太阳、地球和月球。其中，太阳能是主要能量来源。宇宙形成之初，就有了核能和机械能，这些能量在不停

地转化，所以才让这个世界精彩纷呈。

# 2.2　光速是如何测量的？

——伽利略、罗默、迈克尔孙测光速方法

大家都知道：光的传播速度非常快，一秒钟就能走30万公里，一秒钟就可以绕地球七圈半。这么快的速度，人类是如何测量的呢？




## 一、伽利略的测量

在古希腊时代，对于光速的数量级，人们并不是很清楚。一些科学家，比如亚里士多德，甚至认为光速是无限大的。更好玩的是，有人认为，光是从眼睛中发射出来的，我们一睁眼就能看到遥远的星星，所以光速一定是无限大的。

文艺复兴之后，近代科学的先驱伽利略做了第一个测量光速的实验，当时是1638年。



伽利略和他的助手站在两个相隔较远的山头上，每个人手里拿着一盏灯。伽利略首先遮住灯，助手看到伽利略遮住灯之后，立刻遮住自己的灯。伽利略的设想是测量从遮住灯到看到助手遮住灯相差的时间，这段时间内，光刚好在两人之间传播了一个来回，这样就可以测出光速了。

然而，光速如此之快，以至于这个实验根本不可能测出光速。如果不计两人的反应时间和遮住灯的时间，光传播这段距离的时间只需要几微秒，以当时的条件无法完成测量。伽利略也承认，他没有通过这个实验测出光速，也没有判断出光速是有限的还是无限的。不过，伽利略说："即便光速是有限的，也一定快到不可思议。"

# 二、利用木星测光速

真正意义上的光速测量是从丹麦天文学家奥勒·罗默开始的。

1610年，伽利略利用自己改进的望远镜发现了木星的四颗卫星，其中，木卫一最靠近木星，每42.5小时绕木星一圈。而且，木卫一的轨道平面非常接近木星绕太阳公转的轨道，所以，有时候木卫一会转到木星背面，因为太阳的光无法照射到木卫一，地球上的人就看不到这颗卫星了，称为木卫一蚀。

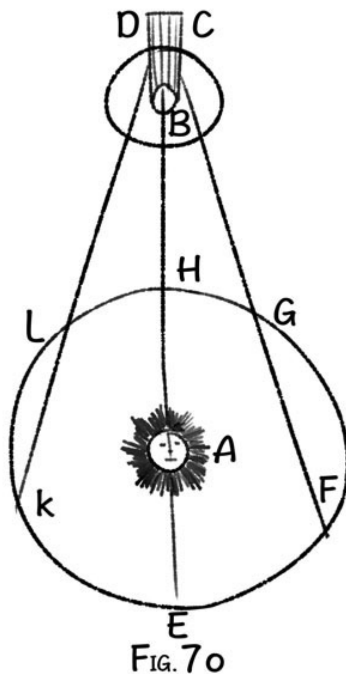


FIG. 70

我们来看一个示意图，地球绕着太阳$A$ 在圆轨道$FGLK$ 上逆时针运动，木卫一绕着木星$B$ 也在逆时针运动。木星背后$CD$ 之间是木星的阴影区，如果木卫一进入这部分阴影，太阳光照射不到它，人们就无法看到它。也就是说，当木卫一到达$C$ 点时就会消失，称为“消踪”，如果木卫一从阴影出来，就能够被人们观察到，也就是木卫一到达$D$ 点时就会出现，称为“现踪”。

罗默就是利用这个现象测量光速的。

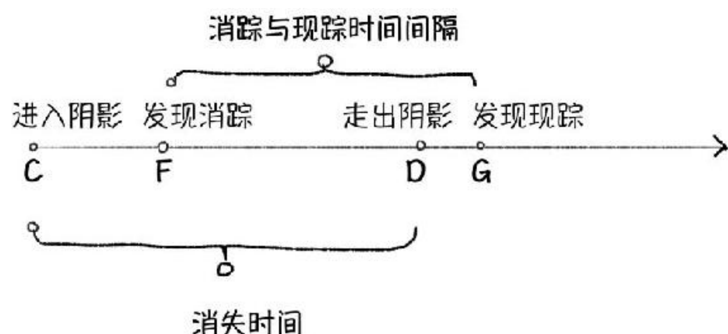
首先，我们研究地球靠近木星时发生的消踪和现踪现象。

当木卫一到达$C$ 点时进入阴影，这个现象的光需要传播一段距离才能到达地球。假设光从$C$ 传播到地球时，地球位于$F$ 点，那么人们观察到消踪现象就比木卫一进入阴影时间晚了一些，这段时间等于$CF$ 长度与光速之比。

当木卫一到达$D$ 点时走出阴影，重新反射太阳光。这个现象也需要一段时间才能到达地球。由于地球在运动，假设这束光到达地球时地球位于$G$ 点，那么，人们观察到现踪现象也比木卫一走出阴影时间晚了一

些，这段时间等于$DG$ 长度与光速之比。

但是，由于$CF$ ，比$DG$ 长，所以消踪现象延迟比现踪现象延迟多一些，即晚发现消踪，早发现现踪。消踪与现踪的时间间隔比木卫一在阴影中的时间要短。我们可以用一个线段图表示这个关系。




同样，我们可以讨论地球选离木星时的消踪和现踪现象。

如果地球到达$L$ 发现木星消踪，到达$K$ 发现木星现踪，由于地球在远离木星，所以$LC$ 的长度小于$KD$ 的长度，早发现消踪，晚发现现踪，人们观察到消踪与现踪的时间间隔就会比木卫一实际在木星阴影中的时间要长。




1671年到1673年，罗默进行了多次观测，并且得出在地球远离木星时，消踪、现踪时间差比靠近时长了7分钟，并得出了光的速度在$10^8$ m/s量级的结论。

牛顿和惠更斯这两位科学巨匠虽然在光到底是粒子还是波的问题上

争执不休，但是在光速测量上，都支持了罗默的方法。牛顿还测量了光从太阳发射到地球需要8分钟的时间，也就是说，我们看到的太阳是8分钟以前的太阳。

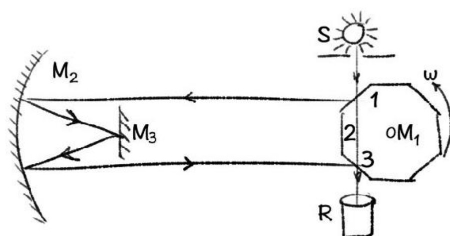



# 三、迈克尔孙和傅科实验

200年之后，第一个把光速测量精度大幅提高的人是美国物理学家迈克尔孙。


迈克尔孙

1877年到1879年，迈克尔孙改进了傅科发明的旋转镜，示意图如下：




迈克尔孙在相隔较远的两处分别放置八面镜$M_1$和反射装置$M_2$、$M_3$，让一束光经过八面镜中的镜面1反射后发出，再通过$M_2$和$M_3$反射回八面镜，经过镜面3反射后进入观察目镜。只有在如图所示的位置时，观察目镜处才会有光。如果八面镜转动一点，经过镜面1反射的光就无法照射到$M_2$，观察目镜上就看不到光了。

如果让八面镜旋转起来，并且角速度逐渐增大，会发现某个角速度下又可以从观察目镜中看到光了。这是因为镜面1刚好倾斜45°角时，光线经过镜面1反射到达$M_2$，再返回八面镜时，八面镜刚好转动一格（1/8周期），于是镜面2刚好跑到图中镜面3的位置，将光线反射进入观察目镜。由于视觉暂留现象，观察目镜中就好像一直可以看到光。

假设左右两套装置相距为$L$，当八面镜转动周期为$T$时，可以从观察镜中看到光。由于$L$远远大于其他部分的长度，所以光从界面1反射到左侧，再回到八面镜走过的距离近似为：

$$S = 2L$$

根据刚才的分析，光来回运动一次，八面镜刚好走过1格，时间为：

$$t = T/8$$

因此光的速度为$v=\frac{s}{t}=\frac{2L}{\frac{T}{8}}=\frac{16L}{T}$。

根据这个原理，迈克尔孙测出光的速度为299853±60km/s，与我

们今天测量的更加精确的值非常接近。

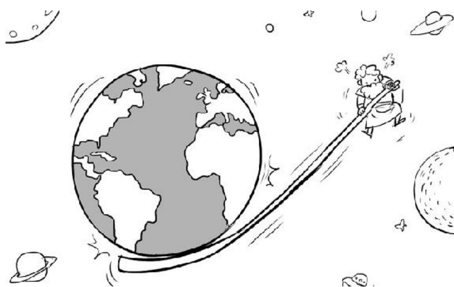



现在，人们使用更加精确的方法测出光在真空中的速度为299792458m/s，并且利用光速来定义“米”的概念。1米就等于光在真空中传播1/299792458秒内传播的距离。如果距离非常大，人们就使用光年的概念：1光年等于光在一年时间里走过的距离，大约$9.5\times10^{15}$ m。我们之所以能看到几百万光年之外的恒星，那是因为那些恒星早在几百万年前就开始发光了，直到今天，它们发出的光才到达地球。

# 2.3　阿基米德能撬起地球吗？

——地球半径和质量的测量

我们小时候就知道阿基米德的名言：给我一个支点，我可以撬起地球。说的是利用一根杠杆可以省力。



阿基米德真的能够撬起地球吗？要做出判断，首先要知道地球的质量，而要测量地球的质量，首先要测出地球的半径。

## 一、测出地球半径的人

人们在很早的时候就知道了地球是球体。最早的学霸毕达哥拉斯第一个提出了地球的概念，而亚里士多德总结了证明地球是球体的三种方法：

1．越往北走，北极星越高；越往南走，北极星越低；

2．远航的船只先露出桅杆顶，慢慢露出船身；

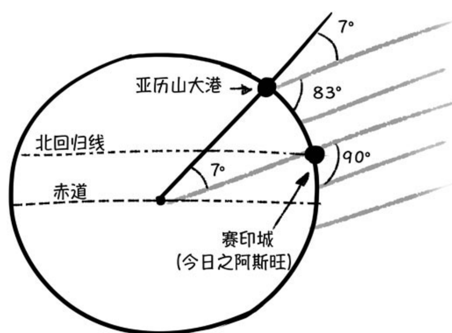
3．在月食的时候，地球投到月球上的形状为圆形。

既然地球是球体，那如何测量地球的半径呢？古希腊的埃拉托斯特尼第一个测量了地球的半径。

他的测量方法是这样的：夏至日，太阳光直射北回归线。而埃及的

城市阿斯旺刚好在北回归线附近，所以夏至日的正午，太阳光会垂直于阿斯旺的水平面，射入阿斯旺的一口深井中。

与此同时，阿斯旺北方的城市亚历山大，太阳光并不直射地面。他通过测量此时亚历山大城中一个石塔的高和影子长度的关系，得到了此时太阳光与垂直地面方向的夹角，大约为7°。




由于太阳到地球的距离远远大于地球的半径，因此太阳光到达地球时接近于平行光。从上图中的几何关系可以看出：亚历山大和阿斯旺与地心连线的夹角就是7°，所以两座城市之间的距离大约是地球圆周长的7/360。通过测量两座城市之间的距离，就得到了地球的周长和半径。如今我们知道，地球赤道的周长大约40000km，而半径大约6400km。

## 二、测出地球质量的人



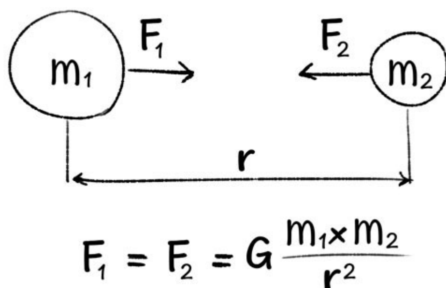

虽然早在两千多年前，地球半径就被测量出来了，但是测量出地球

质量却是十八世纪的事了，也就是三百多年前。

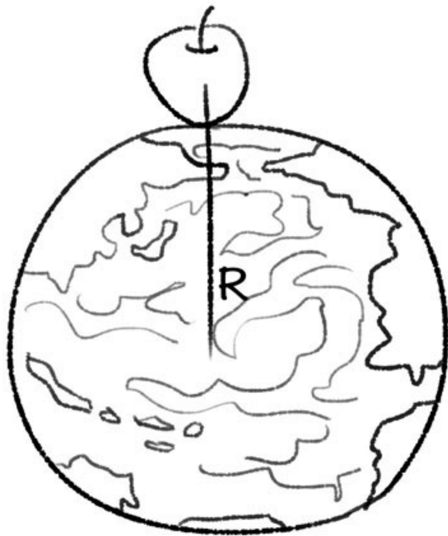
牛顿为了解释苹果为什么能落地，提出了万有引力定律：自然界的任意两个物体之间都相互吸引，引力大小与二者质量的乘积成正比，与距离的平方成反比。写成公式就是：

$$F = G\frac{m_1 m_2}{r^2}$$

其中，$F$ 是引力，$G$ 是万有引力常数，$m_1$、$m_2$ 分别是两个物体的质量，$r$ 是二者的距离。




如果两个物体的尺寸远远小于它们之间的距离，就可以把物体当作点来处理。但是如果物体距离比较近，那么二者的距离究竟从什么地方开始计算，就比较复杂了。但是，如果是质量分布均匀的球体，二者之间的万有引力还是比较好算的，那就是把它们球心的距离代入表达式中的$r$ 即可。




比如，地球上有一个苹果。苹果相比于地球半径很小，所以可以把苹果看作一个点。此时，苹果与地心之间的距离就是地球半径$R$，设地球质量为$M$，苹果质量为$m$，二者之间的万有引力就是：$F=G\frac{Mm}{R^2}$。

这个力就是地球对物体的吸引力，它接近于物体的重力，在这里，我们姑且认为它就等于物体的重力。人们把重力与物体质量的比称为“重力加速度”：$g=\frac{F}{m}=G\frac{M}{R^2}$。

这样，我们就可以得到$GM=gR^2$。这个公式就称为“黄金公式”。

古希腊时代人们就测量出了地球半径$R=6400$千米，牛顿之后，人们又测量出了重力加速度$g=9.8\text{N/kg}$，所以，只需要测量出万有引力常数，就可以知道地球的质量了。

牛顿在1687年的巨著《自然哲学的数学原理》中完整地提出了万有引力定律，但是限于实验条件，牛顿自己并没有测量出这个量。直到100多年之后，英国科学家卡文迪许才通过精巧的扭秤实验测量出了$G$的数值，$G=6.67\times10^{-11}\ \text{Nm}^2/\text{kg}^2$。

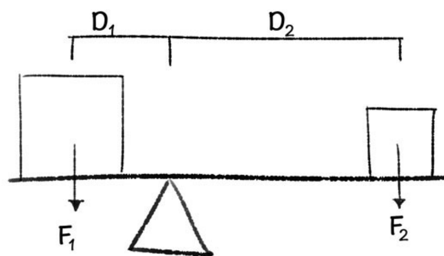
两个质量为1kg的球相距1m时，引力只有$6.67\times10^{-11}$ N，这么小的引力，怪不得牛顿没有测出来。通过以上步骤，人们终于可以计算地球质量了，大约是$6\times10^{24}$ kg。卡文迪许测量了万有引力常数，得以计算

地球质量，所以人们称卡文迪许为“测出地球质量的人”。为了纪念卡文迪许，英国剑桥大学物理系实验室被命名为“卡文迪许实验室”，这也是目前世界上最顶尖的实验室之一。



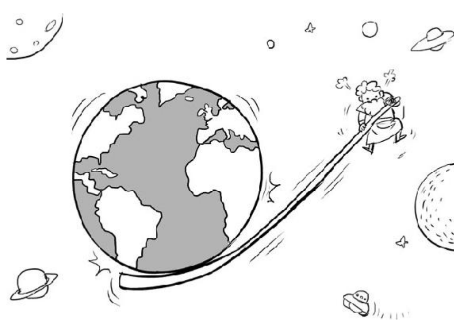

# 三、阿基米德能撬起地球吗？

现在我们终于可以讨论撬地球的问题了！我们知道，阿基米德的时代，人们还没有理解引力的概念。我们姑且认为阿基米德是要在地球上撬起一个与地球相同质量的物体，那么他是否能做得到呢？




根据阿基米德发现的杠杆原理：一个杠杆要平衡，两段施加的力与力臂的乘积应该相等，即$F_1 D_1 = F_2 D_2$，这样一来，如果想用小力去撬动大物体，就需要小力的力臂远远大于大物体重力的力臂。

假设阿基米德有100kg，而地球质量为$6 \times 10^{24}$ kg，如果阿基米德要撬动地球，力臂就需要是地球那一段力臂长度的$6 \times 10^{22}$倍。



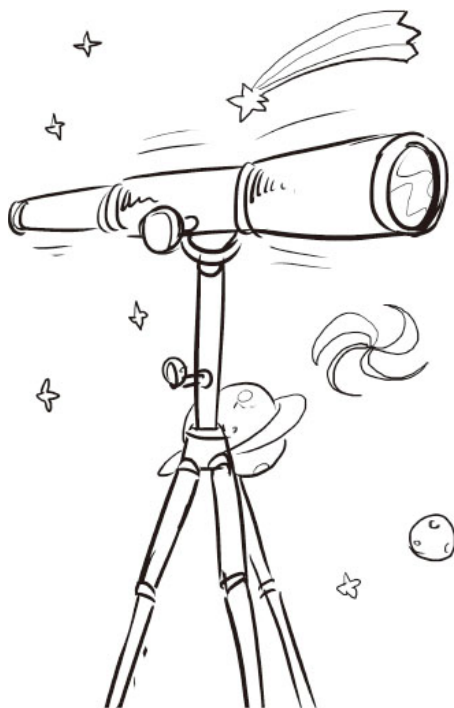
如果阿基米德要把地球撬起1cm，那么根据杠杆臂长的比例关系，阿基米德一端所需要下降的距离就是$6\times10^{20}$ m，大约相当于6万光年。也就是说，阿基米德想凭借自身重力撬起地球的话，即使一切实验设备都准备好了，而且他能够以光速运动，他也需要6万年的时间才能将地球撬起1cm。显然，这是不可能的。

阿基米德的豪言壮语点破了杠杆原理，但是却忽视了地球与人在质量上的巨大差别。

# 2.4　天体之间的距离到底有多远？

——视差法、开普勒定律和金星凌日

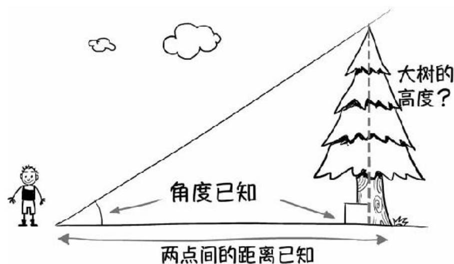
我们晚上看天空，会看到美丽的星星。除了太阳系内部的几颗行星外，大部分肉眼可见的星星都是其他星系的恒星，这些恒星距离我们非常遥远，就算是跑得最快的光到达我们，也需要许多许多年的时间。那么，我们如何测量这些星球到地球的距离呢？



## 一、视差法和秒差距

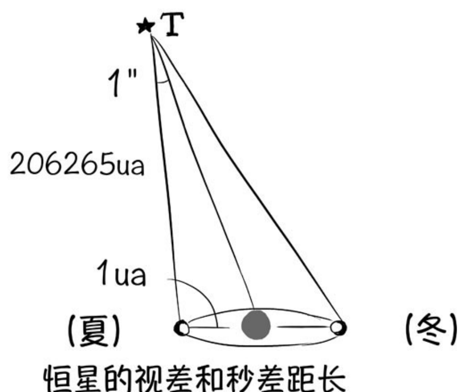
测量恒星的距离，最基础的方法是三角视差法。

我们不妨先从一个简单的例子说起。假如有一棵树非常高，那么我们如何才能测量出这棵树的高度呢？



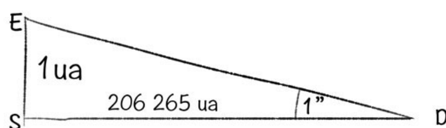

首先我们观察树根和树梢，得出两个观察方向，并且测量它们的夹角。然后我们再测量出观察点和树之间的水平距离，根据三角形的知识，就可以求出树的高度。

三角视差法基本原理与之类似。由于地球绕着太阳旋转，一年中的不同时刻，从地球上观察某个遥远的星球，视线的方向是不同的，我们可以在冬天和夏天记录观察星球时的视线方向，并且测量两个方向的夹角。



恒星的视差和秒差距长

我们知道，一个圆周角为360°，每度又可以分为60′，每分又可以分为60″，于是1″就等于1/1296000圆周，是一个非常小的角度。假如冬天和夏天观察同一颗恒星时，观测方向夹角为2″那么恒星与地球的连线和恒星与太阳连线的夹角就约等于1″，此时我们就称恒星距离地球为一个秒差距（pc）。




我们再把日地距离写作一个天文单位ua，把太阳（$S$ ）地球（$E$ ）

和该天体（$D$　）画成一个三角形，根据三角形的关系可以计算出地球和天体之间的距离$SD$　为一个秒差距，大约是1pc = 206265ua，也就是接近于20万个天文单位。

根据这种方法，人们测量了距离地球最近的恒星——比邻星，它到地球的距离为1.3pc，大约相当于27万个天文单位，银河系中心到地球大约8000秒pc，大约相当于16亿个天文单位。

# 二、开普勒三定律

可是，为了测量出具体的数值，我们还必须测量一个天文单位，也就是地球到太阳的平均距离到底是多少。这个距离又是如何测量的呢？

也许有同学说：我们可以发射一束激光到太阳上，等它反射回来，再测量时间差。这种方法是不行的，因为日地距离太遥远了，我们发射的激光很难到达太阳。就算激光到达了太阳，反射光也会淹没在巨大的太阳辐射光中，没法分辨。

为了了解日地距离的测量方法，我们首先从一个天文学家说起。



开普勒是十七世纪德国的天文学家和数学家。当时普鲁士皇帝鲁道夫二世的御用天文学家是从丹麦来的第谷，而开普勒是第谷的助手。第谷死后，开普勒致力研究第谷留下的海量天文观测数据，写成了巨著《新天文学》。

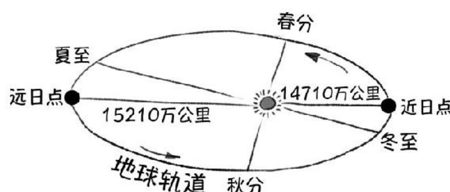
在《新天文学》以及相关著作中，开普勒提出了著名的行星运动三大定律：

1．行星绕太阳做椭圆轨道运动，太阳在椭圆的一个焦点上；

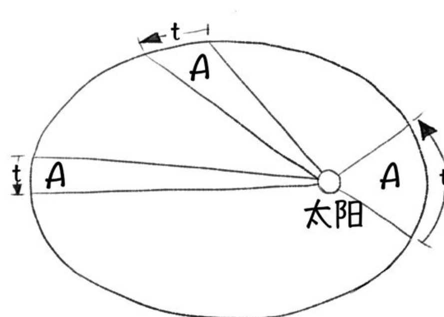
2．行星与太阳的连线在相等时间内扫过相等的面积；

3．行星轨道半长轴三次方与周期二次方的比值是常数。

通过开普勒的研究，人类第一次认识到行星运动轨道是椭圆，而不是圆，因此行星运动时存在“近日点”和“远日点”。地球的近日点是在每年的1月初，远日点是在每年的7月初。不过，地球轨道近日点和远日点与太阳的距离相差不大，地球的轨道还是接近于圆的。




开普勒第二定律是说：如果把行星与太阳做连线，并且终过一段固定的时间，无论行星在何处，这个连线会扫过相等的面积。因此，为了保证面积相等，任何一个星球在近日点处速度都快一些，而在远日点处速度都慢一些。




开普勒第三定律是说：太阳系的行星，轨道半径不同，从小到大依次是水星、金星、地球、火星、木星、土星等。它们的周期也各不相同，而且轨道半径小的周期也小，轨道半径大的周期也大。

| 行星 | 平均轨道半径R/m | 周期T/s |
|---|---|---|
| 水星 | $5.79\times10^{10}$ | $7.60\times10^{6}$ |
| 金星 | $1.08\times10^{11}$ | $1.94\times10^{7}$ |
| 地球 | $1.49\times10^{11}$ | $3.16\times10^{7}$ |
| 火星 | $2.28\times10^{11}$ | $5.94\times10^{7}$ |
| 木星 | $7.78\times10^{11}$ | $3.78\times10^{8}$ |
| 土星 | $1.42\times10^{11}$ | $9.30\times10^{8}$ |
| 天王星 | $2.87\times10^{11}$ | $2.66\times10^{9}$ |
| 海王星 | $4.50\times10^{11}$ | $5.20\times10^{9}$ |

为了简单起见，我们把行星轨道当成圆形来处理。开普勒发现：如果行星的轨道半径三次方与周期平方做比，那么太阳系的几颗行星的这个比值都是相同的。

开普勒是从大量的天文数据中通过拟合和猜想得到上述结论的，但是他并没有解释这是为什么。随后，科学巨匠牛顿受到开普勒三定律的启发，提出了万有引力定律，成功地解释了开普勒三定律的物理内涵。通过开普勒三定律，我们就可以测量日地距离了。

# 三、哈雷和金星凌日



1678年，年仅22岁的天文学家哈雷提出：可以通过金星凌日的办法测量日地距离。这个哈雷就是著名的哈雷彗星的哈雷。

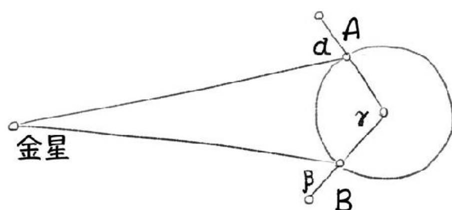
我们知道，金星轨道比地球轨道小，称为“内地行星”。有时候，金

星会经过地球和太阳的连线，称为“金星凌日”，此时在地球上观察，金星像一个黑点一样扫过太阳。

由于地球和金星围绕太阳公转的周期$T_1$ 和$T_2$ 可以通过观测得到，因此根据开普勒第三定律，地球轨道半径$r_1$ 和金星轨道半径$r_2$ 就满足方程：

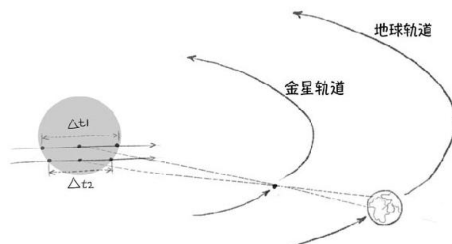
$$\frac{r_1^3}{T_1^2}=\frac{r_2^3}{T_2^2} \qquad (1)$$

而且，金星凌日时，我们还可以通过三角测量法测量地球到金星的距离。我们可以把这个原理简化如下：




在地球上两个地点$A$ 和$B$ 分别观测金星，通过测量金星方向与垂直地面方向的夹角$\alpha$ 和$\beta$ 以及$AB$ 两点对应的地心角γ，再加上人们已经知道的地球半径$R$，就可以通过几何方法计算出金星到地球的距离$d$。而这个距离刚好就是地球轨道半径与金星轨道半径之差。

$$d=r_1-r_2 \qquad (2)$$




联系这两个方程（1）和（2），就可以得到日地距离，也就是地球的轨道半径$r_1$，这就是一个天文单位ua。

遗憾的是，由于金星与地球的轨道并不完全重合，金星凌日的周期比较复杂。金星凌日的时间间隔分别是8年、105.5年、8年、121.5年，

每243年循环一次。也就是说，有的人一生中会遇见两次金星凌日，有的人一生中一次都没有。

哈雷提出这种测量方法后，下一次金星凌日是在83年之后，哈雷知道自己无法亲眼见证这个时刻了，但是人们一直在等待这个时刻。

1761年，人类第一次使用金星凌日测量日地距离，但是很遗憾，没有获得很好的数据。经过精心的准备，8年之后的1769年，英国的科学家在库克船长的带领下，去太平洋上测量金星凌日。当时英法七年战争刚刚结束，英法还处于对峙状态，法国政府特地要求海军不能攻击库克船长的船队。




在当时，航海是一件非常艰苦的事，库克的船在经过了8个月的航行之后，终于到达了目的地——塔希提岛。此时已经有5名船员病死，还有1名船员受不了压力而跳海自杀。1769年6月3日，科学家们终于如愿以偿地观测到了金星凌日。

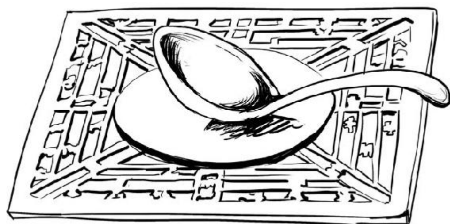
1771年，法国天文学家拉朗德（Lalande）根据这次珍贵的观测资料，首次算出了地球与太阳间的距离大约为1.5亿公里，并命名为一个天文单位ua。人们根据这个数字，推算出了各天体到地球的距离。

在我们教科书上一个简单的数字，都是要经历一代又一代科学家数百年的努力才能得到。我们不得不惊叹于科学的伟大和科学家们孜孜以求的精神。

# 2.5　指南针为啥能指南？

——磁场的形成

指南针是我国的四大发明之一。东汉王充《论衡》中说“司南之杓，投之于地，其柢指南”。是早期对司南比较清楚的描述。考古学家根据古代文献的描述复原了司南，而这是不是指南针的雏形，学术界还有很大的争议。

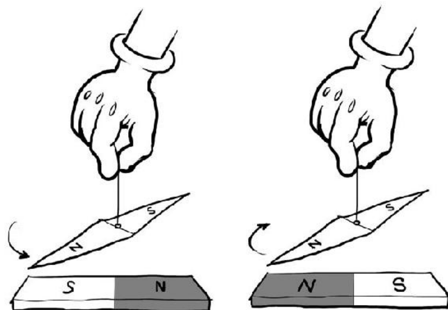


但是没有争议的是，在宋元时期，指南针就已经开始应用于航海等人类活动中。指南针的基本原理就是一个可以自由旋转的小磁针，在地磁场的作用下，小磁针一端指南，一端指北。指南的一端就被称为“南极”（S极），指北的一端就被称为“北极”（N极）。一般我们把指北的一端涂成红色，所以也有人把指南针叫作指北针。

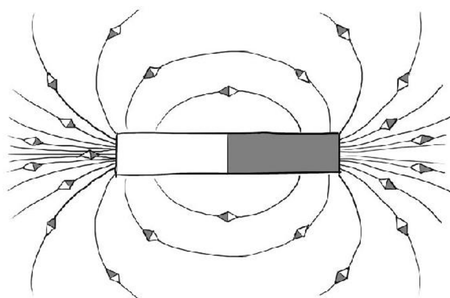
## 一、指南针为什么指南北？

指南针为什么能指南北呢？这是因为地球具有磁场。

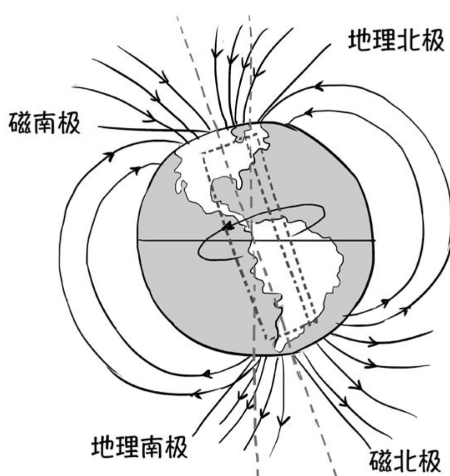
我们知道，磁铁有两极，N极和S极，而且同名磁极相互排斥，异名磁极相互吸引。也就是说，我们用一个磁铁的N极靠近另一个磁铁的S极，就会发现二者相互吸引。如果用一个磁铁的N极靠近另一个磁铁的N极，就会发现二者相互排斥。




为了理解磁铁间的这种作用，人们引入了磁场的概念。磁场是存在于磁体周围一种看不见摸不着的物质，这种物质的作用是对磁铁有力的作用。也就是说，一个磁铁会在周围空间产生磁场，而这个磁场就会对另一个磁铁有力的作用。如果在磁铁周围放置一堆小磁针，那么小磁针的N极（下图中深色部分）指向就会与磁场方向相同。也就是说，在磁体外部，磁场是从磁体的N极指向磁体的S极。



人们根据指南针在地球表面可以指南北的特点，推断出地球是具有磁场的，这就被称为“地磁场”。地球的磁场与条形磁铁的磁场非常像。根据小磁针N极向北指的特点，人们分析出地磁场的的方向是从南向北的，这就得出了地磁场的N极其实在地理南极附近，而地磁场的S极在地理北极附近，地磁南北极与地理南北极是相反的。




而且，地磁南北极与地理南北极也并不是完全重合的，而是存在着一个夹角，称为“地磁偏角”，中国的科学家沈括在《梦溪笔谈》中写道：“方家以磁石磨针锋，则能指南，然常微偏东，不全南也。”是世界上最早关于磁偏角的记录。现代指南针都通过技术手段修正了这个偏角。

# 二、奥斯特实验：电流产生磁场

那么，地球为什么能够产生磁场呢？这个问题更加普遍的问法是：磁场到底是怎么产生的？

最初，人们认为电与磁是完全不相关的现象。但是随着人类认识自然越来越深入，逐渐发现了电与磁可能是相关的，最典型的就是被闪电劈过的铁矿石可能具有磁性。人们想到，也许电可以产生磁。

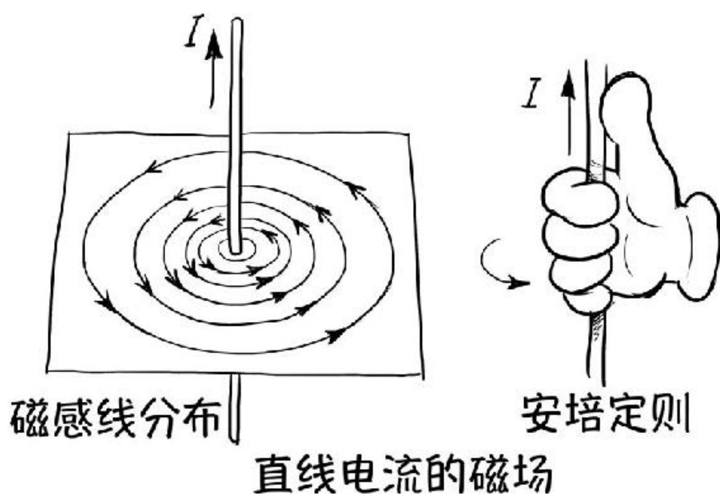
丹麦物理学家奥斯特第一个发现了电流与磁场之间的联系。



1806年，他应聘哥本哈根大学教授，每个月他都会给学生准备一节特别的课程，用来介绍科学界的最新成果。在有一次讲课时，他发现，当用通电导线靠近指南针时，指南针发生了转动。他和他的学生共同见证了这一历史时刻，但是他当时并没有对这种现象给出解释。经过几个月的思索，他终于明白了，电流可以产生磁场，而磁场可以对小磁针有力的作用。此时，人们才正式把电和磁联系在一起。

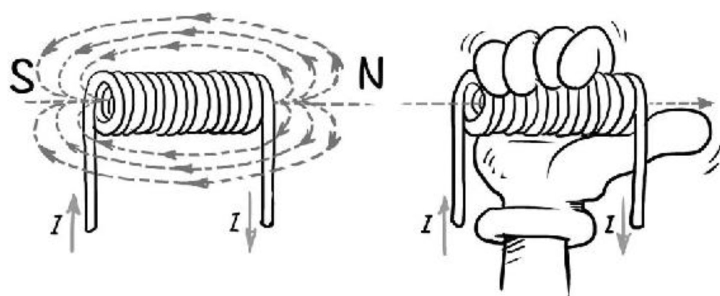
为了纪念奥斯特，丹麦政府把哥本哈根市中心的公园命名为奥斯特公园，里面竖立着奥斯特的雕塑。美国物理教师协会还特别设立了奥斯特奖章，来奖励优秀物理教师。

人们仔细研究了各种形状的通电导线形成的磁场。比如，通电直导线周围的磁场是同心圆环，并且与导线中电流方向符合右手螺旋定则。也就是说，右手握住导线，大拇指指向电流方向，四指的绕向就是磁场方向。



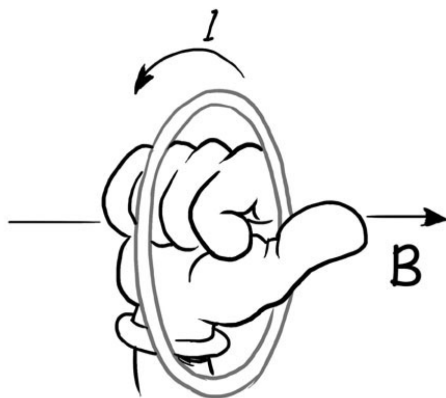

直线电流的磁场

另外还有一种典型情况：通电螺线管。就好像在一根弹簧上通电，所产生的磁场与条形磁铁的磁场很类似。判定方法依然是右手螺旋定则：右手握住螺线管，四指方向与电流方向相同，则大拇指指向螺线管相当于条形磁铁的N极。




通电螺线管的磁场

如果这个螺线管只有一圈，就变成了通电圆环，它的磁场也类似于条形磁铁，判定方法还是右手螺旋定则。



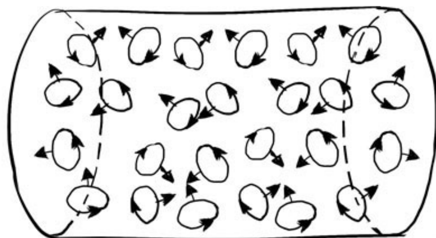

环形电流的磁场

法国物理学家安培最早对这个问题形成了清晰的认识，所以右手螺旋定则也称为“安培定则”。

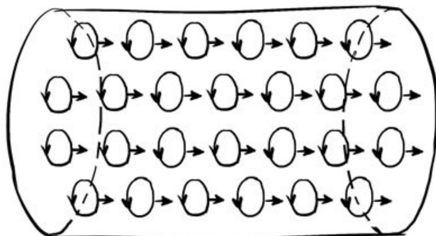



安培详细研究了各种通电导体的磁场，并提出了导线中电流与其产生磁场之间的定量关系，即安培定律。后来，安培的著作《关于电动力学现象之数学理论的回忆录》出版，电动力学作为一个新名词，登上了科学的舞台。

## 三、安培分子电流假说：磁体产生磁场



甲



乙

安培不满足于研究电流的磁场，他进一步思考：既然各种电流都产生相应的磁场，那么永磁体，例如磁铁，它的磁场是如何产生的呢？

安培产生了一个大胆的想法：也许磁体内部也有电流，是这些电流形成了磁体的磁场。这就是著名的安培分子电流假说。

安培认为，在磁体内部可能存在着很小的环形电流，称为“分子电流”。每个分子电流都有N极和S极。如果这些分子电流的取向是杂乱无章的，那么磁场彼此抵消，宏观上就没有磁性。但是如果在外界磁场的作用下，分子电流的取向变得大致相同，那么宏观上就表现出磁场，两端形成磁极。



在安培的时代，人们并不清楚组成物质的原子是什么样的，更不知道原子里面有原子核和电子，所以安培的这种说法只能停留在“假说”阶段。现在的科学界认为，电子围绕原子核的运动和电子的自旋具有磁场，安培分子电流假说是具有一定正确性的。

既然磁体的磁场也是由电流产生的，人们总结出一个结论：一切磁现象都是由电流产生的，或者叫一切磁现象都有电本质。

# 四、地球磁场

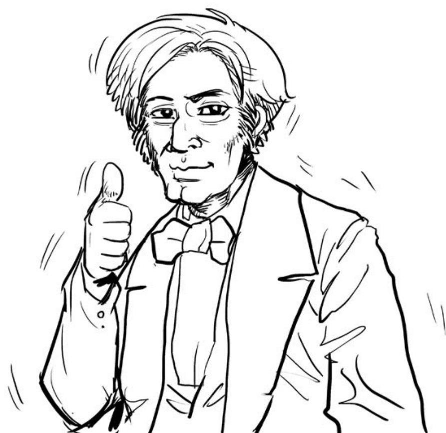
那么，地球的磁场究竟是什么原因产生的呢？对于这个问题，科学界还没有统一的认识。有人认为，在地球内部流动的岩浆会造成电流，产生地磁场；也有人认为在大气中存在电荷，大气运动会造成电流，产生磁场。但无论如何，地球的磁场也一定是由电流产生的。

其实，地球磁场并不是一成不变的，地磁南北两极每时每刻都在缓慢移动，而且在历史上，地球磁场的南北极调换过多次，上一次调换是在78万年前。地磁场对生命有重要意义，例如它可以防止太阳风中的各种射线直接射向地球表面，从而保护地球上的生命。一些生物可能需要磁场进行导航。如果磁场发生巨大变化，那么一定会对生物产生巨大的影响，甚至能造成生物的大灭绝呢。

# 2.6 家用电是怎么产生的？

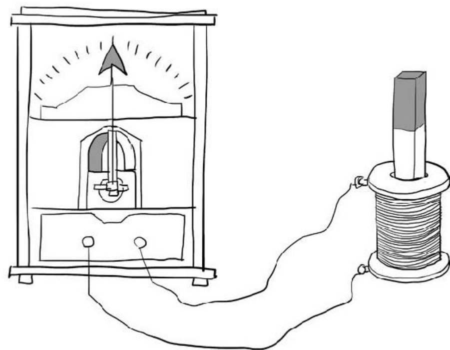
——电磁感应现象

每天我们都离不开电，我们也知道家用电是从发电厂发出来的。但是发电厂是如何发电的呢？是谁第一个发现了这个原理呢？要了解这些，我们首先要从一个人物——法拉第——说起。



## 一、电磁感应现象

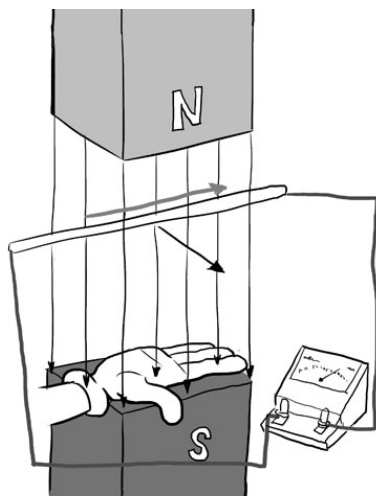
丹麦物理学家奥斯特发现了电流的磁效应之后，英国物理学家法拉第就想到：既然电流能够产生磁场，那么反过来，磁场是否能够产生电流呢？因为在法拉第的时代，人们用电都是使用锌、铜和盐水制作的伏打电池，这种电池制作麻烦、电压小，发出的电不适于普通人使用。但是自然界的磁铁资源非常丰富，如果使用磁铁发电，那么电就能进入千家万户了。



法拉第为了这个理想而进行了艰苦的实验，他最初的想法是将一块磁铁放在螺线管中，期待着电路中产生电流，但是一直没有获得成功。终于，在1831年，法拉第的实验获得了突破，他发现：只有在磁铁插入或者拔出螺线管的过程中，电路中才会产生电流。

我们用现代的方法可以把法拉第的实验等效成上面的情况：用一个螺线管连接电流表，将一根磁铁插入螺线管的过程中，电流表就会产生示数。而且，插入的速度越快，电流表指针的偏转就越大。同样，在磁铁拔出螺线管的过程中，电流表指针也会偏转，只是方向相反。但是，如果磁铁在螺线管中保持静止不动，电路中就没有电流产生了。

法拉第终于明白：只有在运动和变化的过程中，磁铁才能够产生电流。于是，法拉第将他的发现总结成五种情况，其中，应用在现代发电机的情况是：在磁场中切割磁感线的导体可以产生电流。



比如，我们将一根导线与电流表相连，并且使得导线向右运动，这样导线就好像刀一样切断了磁感线，电路中就会出现电流。而且我们可以使用右手定则判断电流的方向：如果磁感线穿过右手的手心，大拇指指向导线运动的方向，那么右手四指的方向就是产生的电流方向。

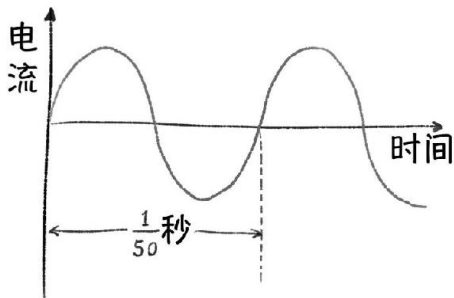
按照这个原理，法拉第制作了早期的发电机，而英国财政大臣对发电机的用途表示怀疑。




果然，现在电已经成了我们生活中必不可少的一部分，谁家用电不交钱呢？

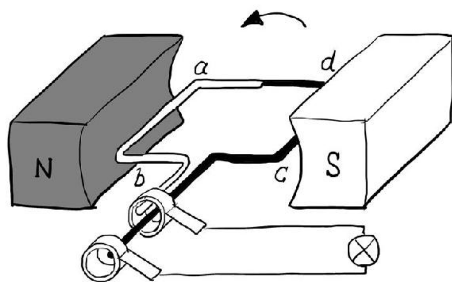
# 二、直流电和交流电

早期的发电机产生的都是直流电。所谓直流电，就是电流方向不发生变化的电流。现在，我们的家用电都是交流电，所谓交流电，就是方向周期性变化的电流。




具体来说，两孔插座中有一根线称为“零线”，零线电压与大地相同，所以触摸零线不会触电。另一根线称为“火线”，火线电压一会儿比大地高，一会儿比大地低。由于电流从高电压流向低电压，所以电流有时候从火线流过用电器，再流回零线，有时候从零线流过用电器再流回火线，每个周期是1/50s，称为50Hz的交流电。交流电与直流电相比，最大的优点就是改变电压十分方便，从而可以进行高压传输以减小损耗。

那么，这种交流电又是如何产生的呢？

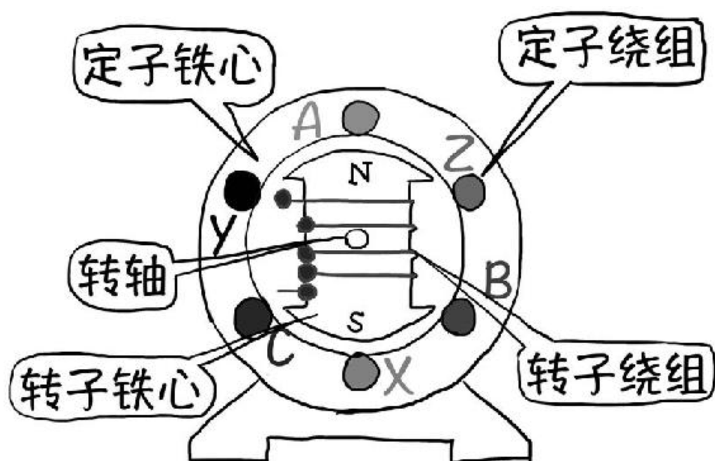


交流发电机的原理

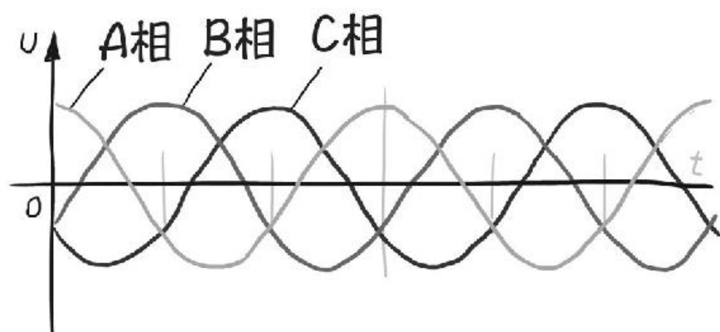
交流电可以通过一个线圈在磁场中转动产生。比如图中这种情况：在线圈旋转时，右侧导线向上运动，根据右手定则，产生的交流电从$c$流向$d$；左侧导线向下运动，产生的交流电从$a$流向$b$，所以整个电路中的电流流向是$c—d—a—b$。在线圈外端通过两个电刷与外部电路相连，于是电流就可以通过电刷从上向下流过灯泡。

半个周期之后，线圈旋转半周，$ab$ 与$cd$ 交换位置，电流方向就会变为$b—a—d—c$，这样一来，电流就会从下向上流过灯泡。于是就形成了交流电。如果线圈匀速转动，就形成了正弦交流电。

也就是说，只要能让线圈与磁铁发生相对转动，就可以产生电流。在现代发电机中，旋转的其实不是线圈，而是磁铁，称为“转子”。而线圈是不动的，称为“定子”。同时，出于工程商的需要，发电机线圈有三组，任意两组线圈都夹60度角。




这样一来，当磁铁在线圈中匀速转动时，三组线圈中分别产生三个正弦交流电，而且每一个都比前一个滞后1/3周期。




这三个交流电会具有共同的零线（中性线）和不同的火线（输出线）。在供电时，如果我们把一根火线和零线接入用电器，也就是家用电220V，如果我们把两根火线接入用电器，就变成了工业用电380V。

那么，我们如何才能使得线圈或磁铁转动呢？这就取决于发电机的种类。水力发电机是靠水的冲击使得涡轮机转动，风力发电机是靠扇叶

使发电机转动，火力发电机是靠燃烧加热水蒸气推动涡轮机。这个过程需要消耗外界的能量。总之，发电机就是把其他形式的能量转化成电能的机器。

发电机被发明之后，各种用电器如雨后春笋般出现，人类自此进入了电气时代。

# 三、没上过学的科学巨匠

我们还要再谈谈法拉第——这个将人类带入电气时代的人。法拉第生于一个铁匠家庭，由于家庭贫困，他只上了两年小学就辍学了，成为了一名订书匠的学徒。不过，这份工作让他接触到了大量的图书，接触到了普通人无法接触到的各种文献，他被科学深深地迷住了。

在书店一位老主顾的帮助下，20岁的法拉第有幸聆听了化学家戴维的演讲。他还把整理好的演讲记录写信邮寄给戴维，并表示了自己愿意为科学献身的想法。

戴维看过法拉第的简历之后，对他说："年轻人，我要告诉你，科学是很辛苦的，而且没有多少回报。"

法拉第回答道："我认为科学本身就是一种回报。"

戴维被感动了，法拉第终于成为了戴维实验室的一名助手，他的科学梦想从戴维实验室开始成为了现实。后来，法拉第在物理、化学方面都取得了惊人的成就，成为了那个时代最伟大的科学家。

法拉第还是一个品德高尚的人。由于自己早年的经历，他非常重视对青年学者的培养，并且鼓励了一批如麦克斯韦一样的科学巨匠。他拒绝了对自己的封爵，还两次拒绝成为皇家科学会会长，并表示不愿安葬在西敏寺——牛顿等人的长眠地。于是，人们将他安葬在其他墓园，却在西敏寺牛顿的墓碑旁树立了他的纪念碑。

顺便说一下，他的入门恩师戴维也是一位著名的化学家，人们认为戴维最大的贡献就是发现了法拉第。

# 2.7　特斯拉和爱迪生谁更厉害？

——交流电与直流电之争

最初法拉第发明的发电机是产生直流电的直流发电机，但是我们今天很多时候使用的都是交流电。那么交流电与直流电相比，有什么优势呢？我们通过讲述历史上一段耐人寻味的“电流大战”来了解一下吧。




## 一、发明大王爱迪生和他的直流发电机

在法拉第发现了电磁感应定律——磁场可以产生电流之后，各种用电器就如雨后春笋般出现了，这里面最重要的一项发明就是电灯泡。有一个人靠改进电灯泡积累了巨额财富，这就是美国发明大王托马斯·爱迪生。

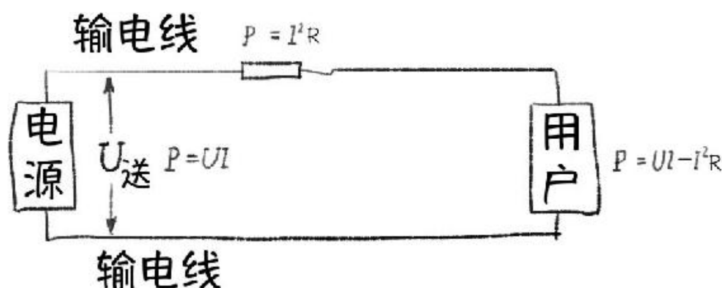


1878年到1880年，爱迪生与助手一起研制电灯泡，并最终找到了可以点亮一千多个小时的灯丝材料。1881年，在巴黎世博会上，爱迪生展出一台重27吨、可供1200只电灯照明的发电设备，从此名扬世界。1882年，爱迪生在纽约金融街旁边创办了自己的直流电力公司，负责向用户提供灯泡等用电设备以及电力供应，并迅速积累了巨额财富。

但是，直流电有一个巨大的缺陷：直流电不能变压。

当发电厂产生电能之后，电能需要通过导线来传输给用户。在用户用电的时候，导线上也会损耗一部分能量。用户的功率$P_{用}$等于发电厂发出来的功率$P$减去导线上损耗的功率$P_{损}$。

发电厂发出来的电功率$P$等于电压$U$与电流$I$的乘积：$P=UI$，电压越大电流越大，传输的功率越大。导线上损耗的功率等于电流平方与电阻$R$的乘积$P_R=I^2R$，电流越大，导线电阻越大，损耗的功率就越大。用户获得的功率是两者之差$P_{用}=UI-I^2R$。



如何才能降低导线上的功率损耗呢？首先人们想到了减小电阻$R$。一般而言，铜的电阻较小，所以导线一般使用铜制作，同时，减小导线长度、增加导线截面积也可以降低导线电阻。但是即便是这样，当用户多、电流大的时候，导线上的损耗依然很大。这么大的热功率不仅浪费电能，更重要的是，会使导线变热，引起火灾。

那么，要进一步降低导线上的能量损耗，就需要减小电流$I$。但是在功率$P$一定的情况下，减小电流就必须提高电压$U$。比如，输电功率同样是100W，如果电压是100V，那么电路中的电流就是1A；如果电路中电压是1000V，那么电流就是0.1A。电流减小了10倍，那么损耗功率就会变为原来的百分之一。

不过，发电厂发出的电压也不能特别高，因为这样的话，用户一端的灯泡全都会因为承受不了高压而爆掉。

那怎么办？爱迪生的回答是："那好吧，我们就每隔1英里建一个发电站，这样导线的电阻就小了。"

显然，这种方式会使得用电的成本大大提高。

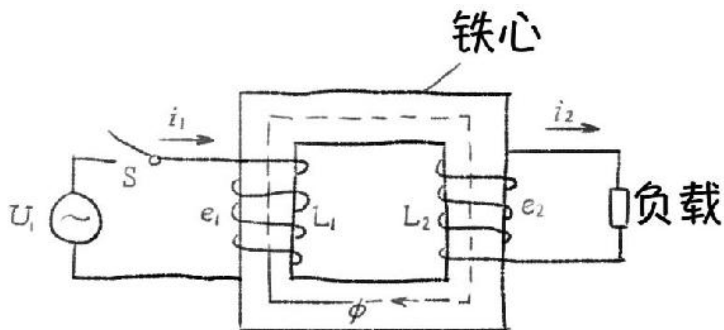
# 二、变压器和高压输电

有没有办法在输电的时候使用高压，到达用户阶段，再使用低压呢？比如，现代的电网基本原理就是下面这样：




首先，发电厂发出的电压通过一个升压变压器变为高压，通过高压电线输电，到达用户一端时，首先通过降压变压器降压，再输送给用户。

为了理解这个过程，我们就需要对变压器知识有所了解。




变压器也是靠电磁感应而工作的。有一个“回”字形的铁心，铁心左端缠绕着一个导线线圈，称为“原线圈”，铁心右端也缠绕着一个线圈，称为“副线圈”。当一个变化电流（交流电）通过原线圈时，由于原线圈电流发生了变化，造成原线圈的磁场也发生变化，这个磁场会通过副线圈，造成副线圈的磁场也发生变化，就好像法拉第实验中的磁铁插入或者拔出线圈一样，副线圈中就会产生电流。而且，根据法拉第电磁感应定律，如果原线圈有$N_1$匝，副线圈有$N_2$匝，原线圈电压$U_1$，副线圈电压$U_2$，那么两个电压之间的关系为：

$$U_1:U_2=N_1:N_2$$

通过调整两个线圈的匝数关系，就可以实现升压和降压了。

但是变压器要工作，必须使用交流电。因为恒定电流通过原线圈时，由于电流不变，磁场也不变，副线圈中就不会产生电磁感应现象。于是，一些科学家就想到：是否应该使用交流电来代替直流电。

# 三、交流电之父：尼古拉·特斯拉

说到交流电，就不得不提到尼古拉·特斯拉。



特斯拉是塞尔维亚的科学家，年轻时就一直思考着交流电和交流发电机的构想，但是一直没有实现。他觉得世界上有一个人可以帮助他实现梦想，这个人就是爱迪生，于是，他在1884年来到纽约，并加入了爱迪生的公司。

特斯拉向爱迪生阐述了自己的想法，但是并没有获得爱迪生的支持，不过爱迪生还是雇佣了特斯拉，并请特斯拉改进自己公司的直流发电机。爱迪生许诺：直流发电机改好之后，特斯拉将获得5万美元的奖励。当时的5万美元不是个小数字，可以在纽约买一所房子。

年轻的特斯拉夜以继日地工作着，每天从早上10点工作到第二天凌晨5点。但是当他工作完成，并向爱迪生要报酬的时候，爱迪生只是大笑，并对他说："特斯拉先生，看来你是不懂美国式的幽默。"



特斯拉一怒之下离开了爱迪生的公司，通过干体力活维持自己的生计。一年多以后，在投资人的资助下，特斯拉终于创办了自己的交流电力公司，并生产各种基于交流电的发电、变压和用电设备。西屋电气公司的老板乔治·豪斯在看了特斯拉的展示之后，购买了特斯拉的所有专利，并以3.5美元每千瓦时向特斯拉支付版权费。如果这个版权延续到今天，那特斯拉积累的版权费会是一个天文数字，因为现在全世界都在使用交流电，2017年，全球发电量24万亿千瓦时，而其中大部分都是交流电。

# 四、电流大战

由于交流电对比直流电有着先天优势，以直流电为主业的爱迪生对此十分恼火，并决定通过各种方式诋毁交流电。爱迪生向公众强调：交流电是危险的，必须禁止。他使用交流电公开处决如大象等大型动物，并把处决过程拍摄成电影给公众播放。他甚至通过自己的关系游说国会，将死刑犯的行刑方式从绞刑改为了电椅。

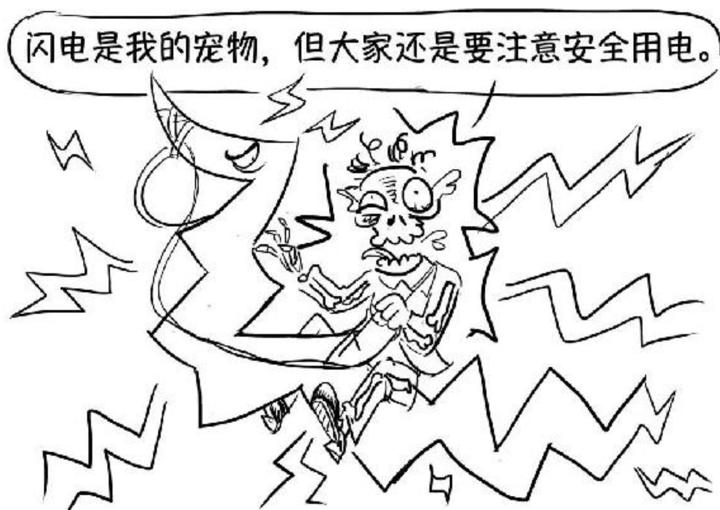
1890年，世界上第一例电椅实验在纽约的监狱展开，由于经验不足，火苗从罪犯的脊柱蹿出，但是罪犯却还活着。

这场交流电和直流电的战斗在1893年芝加哥世博会上进入了白热化。芝加哥世博会是人类历史上第一次使用电灯照明的世博会。此时，爱迪生公司已经改名为通用电气公司，他们的首要任务就是要揽到这个活。他们对政府开价大约100万美元。但是，特斯拉的公司开价只有他们的一半，显然，爱迪生没有拿到这个工程，而特斯拉得到了一个创造历史的机会。

1893年5月1日，十万观众涌入世博园区，夜幕降临，克里夫兰总统按下按钮，整个世博园区的数万盏灯泡被特斯拉的交流电点亮，人类从未见过如此的情景，对那个时代的人来说，这就是未来。

在世博会上，特斯拉为了消除公众对交流电的恐惧，向人们展示了奇妙的实验：他身着礼服，脚穿木鞋，双手接通电路，用身体作为导线通过交流电，全身闪着电火花。而且在那之后，他似乎就迷恋上了闪

电，他经常在放着闪电的实验室看书，这种让人看着都害怕的东西仿佛是特斯拉的宠物一样温顺。




世博会给人们留下了深刻的印象，特斯拉也一度成为当时的风云人物。特斯拉在小的时候有一个梦想——将尼亚加拉水瀑奔腾的水流变为电能。现在就是他实现梦想的时候了。

1896年，特斯拉主持建造了尼亚加拉水电站，尼亚加拉瀑布强大的水流推动了巨大的发电机，提供了4000kW的电能，通过变压器升压到22000V，再通过高压电线输送到360英里之外的纽约，通过降压变压器降压之后，给交流电动机、电灯泡和电车等用电设备供电。在特斯拉的带领下，电能真正走进了千家万户。

# 五、普罗米修斯式的科学家

特斯拉发明了特斯拉线圈，促进了交流电的应用，发明了交流电动机，还有其他数百项专利，有一种说法是：科学界有两个被严重低估的天才，一个是莱昂纳多·达芬奇，另一个就是尼古拉·特斯拉。

他虽然是一个卓越的科学家和工程师，但却不是一个成功的商人。他没有从自己发明的交流电中获得太多收益，而且由于他太醉心于无线输电技术——一种不用导线就可以输电的新技术的研究，而被资本家们

所抛弃。

世界没有认识到这位天才的价值，晚年的特斯拉孤独地住在纽约的小旅馆里，被出租车撞了，却无钱医治，还是以前的东家——西屋电气公司帮他支付了医药费。

在第二次世界大战时，特斯拉曾经向美国政府提了两个建议：天气风暴和死光武器。1943年1月8日，总统安排特斯拉在白宫会面，但是前一天夜里，特斯拉在他住的小旅馆里去世了。纽约市长在电台中发布了讣告：

“昨晚，一位87岁的老人去世了。他去世时一贫如洗，但是却是对这个世界贡献最多的人。如果把他的发明从生活中抽走，那么工厂车间将停止运转，电车将停止运动，我们的城市将陷入黑暗。特斯拉并没有死，他的生命已经融入了我们的时代。”




特斯拉两次和诺贝尔奖失之交臂，第一次是无线电的发明。他最早进行了无线电通信的演示，并获得专利，但是美国专利局却随后撤销了他的专利，转而授予马可尼，并使马可尼获得了1909年的诺贝尔奖。1915年，纽约时报头版刊文：瑞典政府已经决定将本年度的诺贝尔物理学奖颁发给托马斯·爱迪牛和尼古拉·特斯拉。特斯拉也登上了时代周刊的封面。但是，一周之后，诺贝尔奖却授予给了牛津大学的布拉格。传说

这是因为特斯拉拒绝与爱迪生分享诺贝尔奖，而这个谜团至今也没有得到解释。

为了奖励特斯拉的贡献，美国电气工程师协会（现合并为电气电子工程师协会IEEE）决定授予特斯拉他们协会的最高奖章——爱迪生奖章。

显而易见，特斯拉没有去领奖。

# 2.8 SOS是什么意思？

——无线电报的原理

我们经常看到战争片里有个发报员，“嘟嘟嘟”地发电报。也有很多人知道，SOS是求救信号。那么，大家知道电报是什么原理吗？哪些人对无线电报的发明有贡献呢？人们为什么把SOS作为求救信号呢？

S · · ·

O — — —

S · · ·

## 一、麦克斯韦：电磁波的预言者

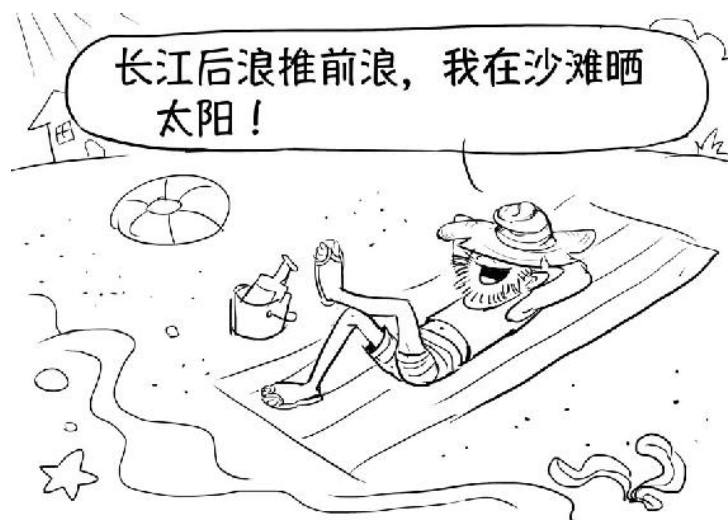


之前我们说过，法拉第发现了电磁感应现象，通过磁场产生电流，并且发明了发电机。可是，为什么磁场会产生电流呢？法拉第并没有想清楚。因为法拉第小的时候家庭贫困，只上过两年小学就辍学了，他的伟大成就都是依靠对科学的热爱和努力钻研的精神而取得的，所以法拉第在数学方面不是很强，总是喜欢用形象的方式表述物理规律，却难以使用数学语言解释自己的伟大发现。

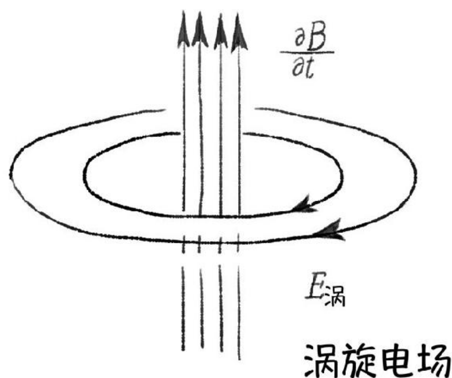
这时，另一个伟大的人物麦克斯韦登场了。

麦克斯韦比法拉第小40岁，在他23岁刚刚从剑桥大学毕业时，读到了法拉第的科学论文，他被法拉第深邃的思想所吸引，并决心用数学去弥补法拉第的不足。一年后，他就发表了第一篇关于电磁学的论文，并同法拉第进行了深入的讨论。

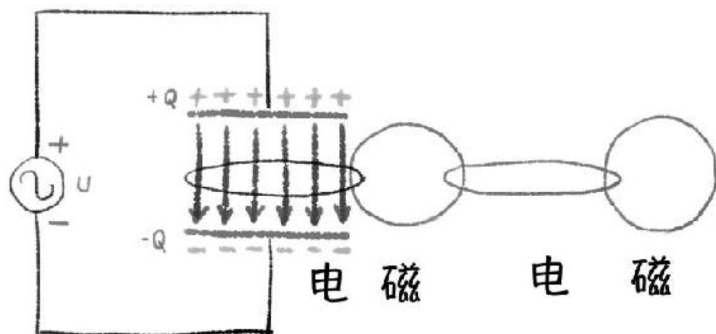
法拉第是一位伟大的科学家，同时也是一个品格高尚的人，他对年轻的学者特别关心。他对麦克斯韦说：你不要只局限于用数学解释我的观点，而要有所创新。在老一辈科学巨匠的鼓励下，麦克斯韦最终成功地提出了麦克斯韦方程组，成为了牛顿和爱因斯坦之间最伟大的物理学家。



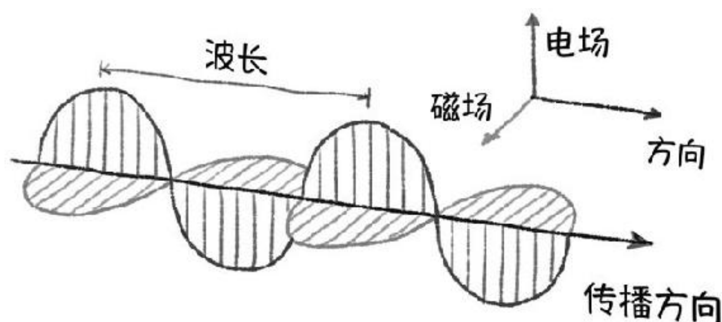

麦克斯韦想：电流的形成是由于电荷的运动，而只有电场能够驱动电荷。他敏锐地感到：变化的磁场并非直接产生电流，而是在周围空间产生了电场，如果在电场附近存在导体，就会形成电流。




麦克斯韦进一步想到，由于电与磁紧密相关，而且变化的磁场能够产生电场，那么变化的电场应该也能产生磁场。比如我们使用交流电连接两个金属板构成的电容器，由于交流电的电压在反复变化，电容器中就会产生变化的电场。




更为神奇的是，如果变化电场产生的磁场依然是变化的，它就会进一步产生电场，如此一来，振荡的电场和磁场可以相互激发，并向远处传播，形成一种类似波动的物质，并命名为电磁波。




麦克斯韦通过计算得到了电磁波的速度，恰好与光速相同，于是麦

克斯韦大胆预言：光就是一种电磁波。遗憾的是，他没有亲手证实自己预言的电磁波，1879年，麦克斯韦在剑桥病逝，年仅48岁。而在那一年，二十世纪最伟大的科学家爱因斯坦刚好出生。

## 二、赫兹：实验证实电磁波

科学的接力棒传到了德国科学家海因里希·鲁道夫·赫兹手中。

赫兹在柏林大学学习时，受到导师亥姆霍兹的指导而研究麦克斯韦电磁理论，并决心用实验证实麦克斯韦的观点。1888年，赫兹设计了一套实验装置，仔细地研究了这种波的波长、频率等信息，得出了这种波的速度等于光速，与麦克斯韦的预言完全一致。至此，电磁波彻底被人们证实了。赫兹实验不仅证实了麦克斯韦的电磁理论，更为无线电、电视和雷达的发展找到了途径。我们今天一刻也离不开的手机就是通过电磁波传输的。




利用无线电传输信号有很多好处，比如真空不能传播机械波，我们在宇宙中喊话，别人是听不见的，但是电磁波却可以在真空中传输，所以宇航员在月球上即使距离再近，也需要使用无线电通信。无线电信号传输速度非常快，如果我们从美国纽约喊一句话，即使无阻碍地传播，北京的人听到这句话也要9个小时之后。但是无线电的传播速度是光

速，几乎可以在一瞬间传遍全球。而且，电信号比机械信号更容易进行放大和信息处理。从电磁波发现的那一天起，人们就一直在研究如何使用无线电进行通信。

## 三、无线电报的发明

在电磁波被发现之前，人们已经在使用有线电报。但是这种方式面临着诸如铺设导线、维护保养等诸多问题，于是研究无线电报被许多科学家和商人提上了日程。对无线电报发明有贡献的科学家有三位：美籍塞尔维亚科学家尼古拉·特斯拉，意大利科学家马可尼，俄罗斯科学家波波夫。

1893年，尼古拉·特斯拉在美国密苏里州圣路易斯首次公开展示了无线电通信，并于1897年在美国获得了无线电技术的专利。然而美国专利局于1904年将其专利权撤销。这一举动据说是因为特斯拉太醉心于他所梦想的通过无线方式传输电能以造福人类的新技术，而忽视了资本家迫切需要无线电报的感受，这些资本家可能有托马斯·爱迪生，安德鲁·卡耐基和摩根等人，于是他们把橄榄枝投到了新的科学家身上，这个人就是意大利人马可尼。



1901年，马可尼发射的无线电信息成功地穿越大西洋，从英格兰传到加拿大的纽芬兰省。1909年，无线电第一次发挥作用，有一艘汽船由于碰撞遭到毁坏，沉入海底，由于无线电的作用，大部分船员获救，同年，马可尼获得诺贝尔奖，马可尼因被称为“无线电之父”而为世界所知。

几乎是在同一时期，俄罗斯科学家波波夫也发明了无线电装置——收音机。他还在收音机上装了天线，这是人类的第一根天线。1896年，波波夫在俄罗斯物理化学协会的年会上，正式用无线电传输了一段信息，信息内容是“海因里希赫兹”，以表示对赫兹的尊重。

究竟是谁第一个发明了无线电，不同的国家有不同的说法。就算是在一个国家，说法也在反复变化，比如美国虽然在1904年撤销了特斯拉的专利，转而授予马可尼，但是在几十年后的1943年，美国最高法院又撤销了马可尼的专利，仍裁定特斯拉为无线电的发明者。无论如何，他们都是伟大的科学家，为我们的世界提供了无限的便利。

# 四、摩尔斯代码

无线电报的基本原理非常简单，就是发报机接通电源，产生一个或长或短的电磁脉冲，电磁波传播到接收端之后，被收报机探测到。最初的无线电信号都是以摩尔斯码的形式进行传输的，这是因为摩尔斯码将英文字母和数字都编码成点和线两个状态，编码简单，容易传输。我们经常在电影中看到战争片中的发报员在那里嘟嘟嘟地按着按钮，这就是在使用摩尔斯电码发电报。

我们熟悉的求救信号SOS，本身并没有任何含义，只是因为它在摩尔斯电码中的表示非常简单，三个点，三条线，再三个点。所以国际无线电报公约组织就把它定为了国际通用的求救信号。

发报员首先将信息转化成摩尔斯代码，通过类似于赫兹实验的发报装置，把长短无线电信号按照一定的规则发送出去。接报员利用同样的装置接收信号，再通过人工方式将信号代码转化成文字信息。显然，这种方式传输信息的效率不高，但是在100年前，这已经是最先进的通信

方式，二十世纪，拍一份电报都要去电报大楼排队，收费都是按字数收，所以都要把内容写得越简单越好。

直到二十世纪末，我国一些地区还在使用电报进行通信。进入二十一世纪之后，随着电话、互联网等通信方式的普及，电报才逐渐退出了民用通信的领域。

现在，摩尔斯代码和无线电报成了一些爱好者的玩具，有人使用摩尔斯代码进行交流，反而感觉非常时尚。

| | | | |
|---|---|---|---|
| A .- | J .--- | S ... | 2 ..--- |
| B -... | K -.- | T - | 3 ...-- |
| C -.-.. | L .-.. | U ..- | 4 ....- |
| D -.. | M -- | V ...- | 5 ..... |
| E . | N -. | W .-- | 6 -.... |
| F ..-.. | O --- | X -..- | 7 --... |
| G -.-. | P .--. | Y -.--- | 8 ---.. |
| H .... | Q --.- | Z --.. | 9 ----. |
| I .. | R .-. | 1 .---- | 0 ----- |

# 2.9　FM和AM啥意思？

——广播信号的发射、传播和接收

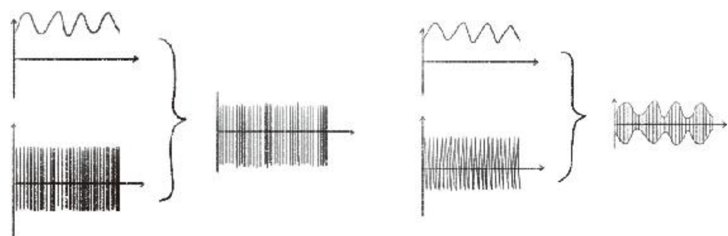
我们经常听到广播里说：FM96.6MHz，AM927kHz。FM和AM到底是什么意思呢？我们又是如何从发射器发射音乐给收音机的呢？

之前我们讲到：利用无线电发射装置可以将信号发射出去，人们根据这个原理制作了无线电报。但是，无线电报传播的是摩尔斯代码，而广播要求传输声音，这又该怎么做呢？

由于无线电必须频率足够大才能够进行有效传输，比如广播频率多数是几百千赫兹到几百兆赫兹，中央人民广播电台经济之声，就是FM96.6MHz，它表示每秒钟电磁波会有9660万个周期。但是我们说话的声波频率低得多，只有几百赫兹。我们如何才能把我们说的话变为无线电信号呢？这就涉及一个概念：调制。

## 一、信号调制

就好像寄快递，我们需要用一个盒子把货物装起来。低频信号是我们所需要传输的，就好像货物；我们再产生一个高频信号，他是我们的载体，就像盒子。我们让高频信号随着低频信号变化，就好像把货物装载在马车上一样，这个过程就称为调制。如果我们让高频信号的频率随着低频信号变化，这个过程就称为调频或者FM。

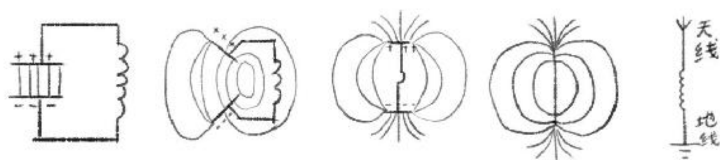


如果让高频信号的振幅随着低频信号变化，这个过程就称为调幅或

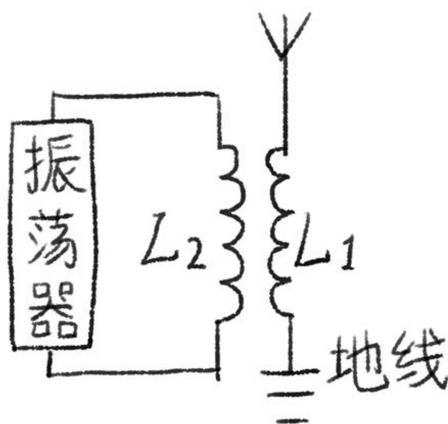
者AM。

## 二、无线电波的发射

调制好的信号就可以进行发射了。一个电容和一个电感构成的电路称为$LC$ 回路，在$LC$ 回路中会产生振荡的电磁场，向外发射电磁波。人们研究发现：电路中的电容越小，发射的无线电频率越高，于是为了发射高频信号，我们把电容的两个板子面积减小。同时，为了让电磁波发射的范围足够大，我们可以把电容器的两个极板一个放置在顶端，一个放置在底端，这就构成了天线。




我们把调制好的电流信号通过电磁感应加载到天线上，天线就可以帮我们把无线电发射出去了。




## 三、无线电波的传播

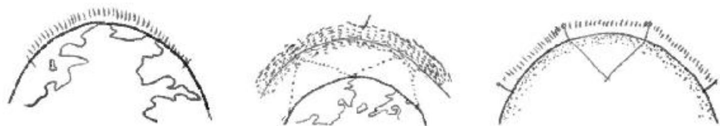
无线电在空中是如何传播的呢？大家注意，由于地球是圆的，如果无线电是沿着直线传播，就很难扩展到很大的范围。人们如何让无线电

传输得更远呢？一共有三种方式。

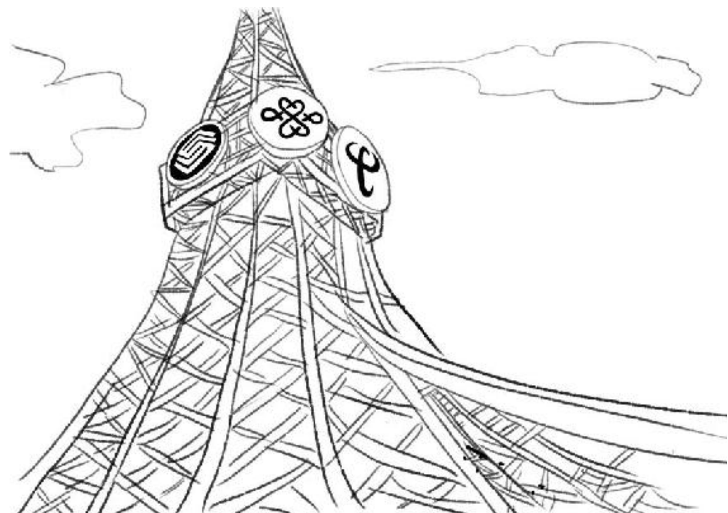
波长比较长的无线电称为长波，长波容易绕过障碍物，这种现象称为衍射，所以长波主要靠地波传播。所谓地波，就是让无线电沿着地面运动，它会自己发生衍射而弯曲，到达接收装置。这种无线电一般用于远程无线通信。

波长短一些的无线电称为中波，中波衍射能力差，不能靠地波传播。但是，在大气中存在着一层特殊的部位，称为电离层，电离层可以反射中波，于是可以利用无线电在地面和电离层之间的反复反射传输无线电，这种方式称为天波。而电离层与天波传输的设想，最早也是特斯拉提出的。电报和广播一般都使用天波传输。

波长比中波短的无线电称为短波和微波。这种波波长很短，难以衍射，所以不能使用地波传输。同时这种波容易穿透电离层，也不能靠天波传输。短波和微波的传输方式一般是直线传播，就是通过高架的铁塔上的中继器将信号一段段接收放大然后继续传输。

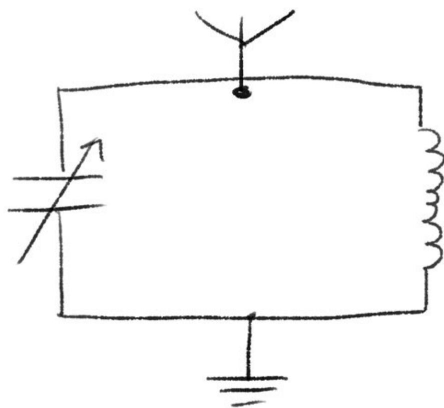


手机无线电信号传播的原理就是直线传播，最开始移动、联通、网通、电信等几大运营商每家都要建设铁塔和基站，十分浪费。最后国家出面协调，由几大运营商共同出资组建了中国铁塔公司，专门用来建设基站塔，实现了资源共享。



由于人造地球卫星的出现，人们更多的把这种信号直接发射到卫星上，然后再通过卫星传输到地面上的另一个地点，这样就节约了基站。

# 四、无线电波的接收



收音机上也有天线，天线是导体，无线电信号遇到导体就会在导体上激发起同样规律的感应电流，只是这个电流比较微弱。不过，接收天线也是$LC$ 回路，它也存在固有频率，如果接收天线的频率与空间中的无线电频率相同，就会产生电磁共振，此时天线上的感应电流最大。

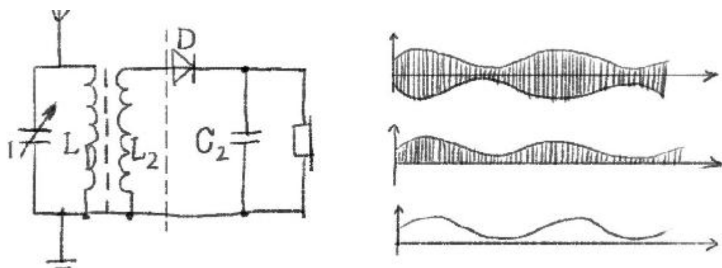
比如，我们想收听FM96.6，就把接收装置的固有频率调整到96.6MHz，此时电路中收到经济之声广播引起的感应电流最大，而其他

电台信号虽然也在电路中有感应电流，但是由于没有共振，所以电流很小。

调整接收装置固有频率的过程其实是调整收音机中的电容大小，这个过程称为调谐，也就是我们生活中所说的搜台或者换台。

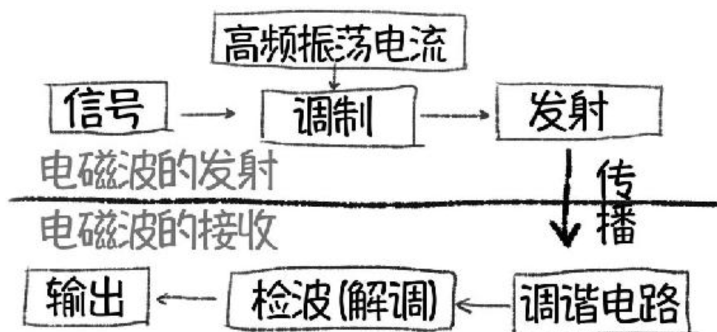
# 五、信号解调

当我们接收到FM96.6MHz的经济之声频道之后，是不是就可以直接把电流通向喇叭了呢？答案是不可以。因为我们接收的信号是经过调制的高频信号，就好像接收了一个快递，首先要把快递拆开一样。我们需要把低频信号从高频信号中检出来，这个过程称为解调。




我们按照与调制相反的方法，把高频信号过滤掉，余下低频信号。再把这个信号放大，通入喇叭。喇叭中的线圈在磁铁对电流作用力下前后振动，带动发声的膜片，我们就可以接收到广播中播放的优美的歌声啦！

总结起来说，广播的过程就是：信号经过调制变为高频信号，通过天线发射出去，经过各种方式到达接收端，经过调谐被接收装置接收，再经过解调就变为了原来的信号。



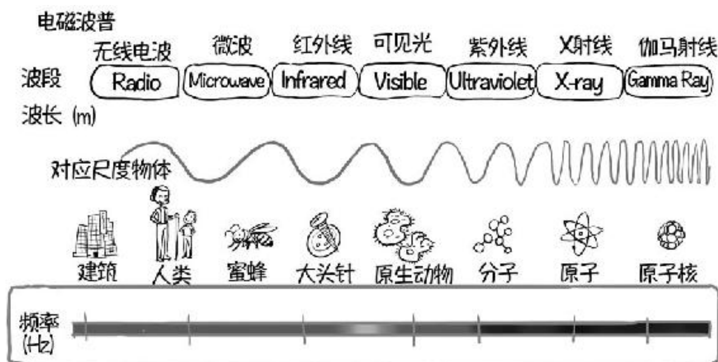

电视、手机等无线电装置的原理是相同的。但是也有一点区别。广播信号一般是可以连续变化的，这种信号称为模拟信号；而手机信号一般采用数字信号，所谓数字信号，就是只有0和1两个状态，与计算机原理相同。利用数字信号，可以方便地实现信息的处理和运算。

# 2.10　世界上第一张X光片是谁拍的？

——电磁波的种类、产生和应用

我们知道，其实光是一种电磁波，通信用的无线电也是电磁波，其实医院里用的X射线，治疗疾病用的伽马射线，也都是电磁波。电磁波的本质相同，都是电场与磁场相互激发，但是为什么它的特点和作用有这么大差别呢？它们都是如何产生的呢？

波，包括电磁波和机械波，都有三个参数：波长$\lambda$，频率$f$和波速$v$，它们三者的关系是波速＝波长×频率，即$v=\lambda f$。对于电磁波而言，在真空中的传播速度$v$和光速$c$一样，即30万km/s。所以波长$\lambda$越大，频率$f$就越小。按照波长和频率的不同，人们把电磁波排列出来，就称为“电磁波谱”。

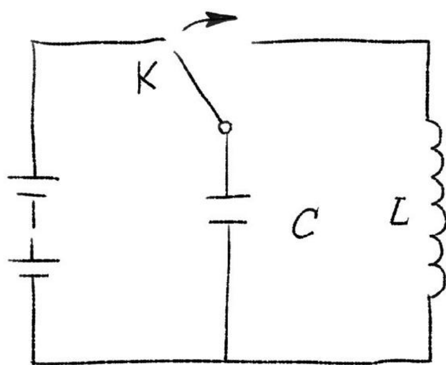



我们按照电磁波产生方式和特点的不同，大致可以把电磁波分为无线电波、光波、X射线、伽马射线四类。

## 一、无线电波

无线电波是波长最长、频率最低的电磁波。它的波长大于1mm，频

率小于300GHz。我们的手机、电视、广播等都使用无线电进行通信。要发射无线电波，就需要有周期性变化的电流，这就需要我们之前说过的$LC$ 回路。



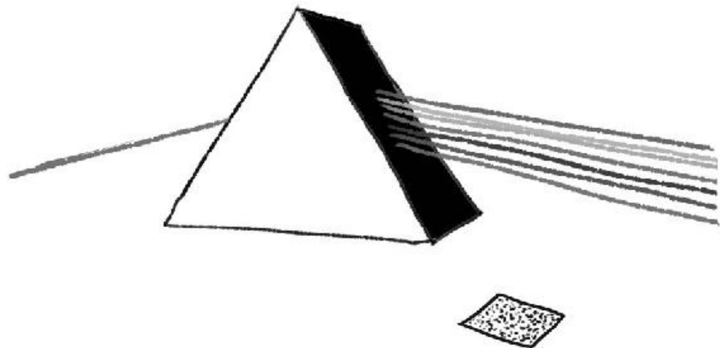

在一个电路中串联一个电容器$C$ 和一个电感$L$ ，由于电容和电感的作用，电路中会产生周期性的电流，对电容器反复充电和放电。根据经典电动力学，此时就会产生变化的电磁场，从而形成电磁波。在收音机、手机里都有这样的装置，只是还需要我们对信号进行调制。

不同的通信装置无线电波的要求不同，比如手机信号使用的无线电波长很短，称为“微波”。而广播信号使用短波或中波。发射信号时天线的长度需要与电磁波波长接近。因此手机天线一般比较短，有时会隐藏在手机内部；收音机的天线比较长，在收音机外部。有些手机具有收音机功能，但是必须插上耳机才能用，这就是因为此时的耳机充当了收音机天线的作用。

# 二、光

波长比无线电短的电磁波称为“光”，又可以分为紫外、可见和红外三个波段。

紫外线波长在5～370nm之间，由于波长比较短，所以肉眼不可见。紫外线是德国物理学家里特发现的。1801年，他在研究太阳光谱时，突然想要了解太阳光分解为七色光后，有没有其他看不见的光存在。



人们当时已知道，氯化银在加热或受到光照时会分解而析出银，析出的银由于颗粒很小而呈黑色。当时他手头正好有一瓶氯化银溶液，于是他用一张纸片蘸了少许氯化银溶液，并把纸片放在白光经棱镜色散后七色光的紫光的外侧。过了一会儿，他果然在纸片上观察到蘸有氯化银部分的纸片变黑了，这说明太阳光经棱镜色散后，在紫光的外侧还存在一种看不见的光线，里特把这种光线称为“紫外线”（UV）。

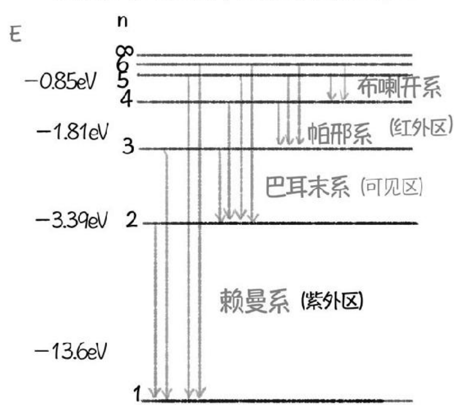
紫外线虽然肉眼不可见，但是对生命是具有重要意义的。紫外线可以杀菌，因此医院里用紫外线灯为一些器材消毒。

同时，紫外线可以促进钙的吸收，太阳光里有紫外线，适当晒太阳对人的身体健康有好处。经常在矿井里工作的工人长期受不到阳光照射，每隔一段时间还需要使用人工的紫外线灯照射。丹麦的医生芬森对阳光治疗疾病问题很有研究，传说他们家有一只猫，有一次猫受伤了，就在阳光下趴着。当太阳移动了位置之后，猫会重新找到阳光继续趴着。从此他就着迷于阳光对健康问题的研究，并获得了1903年诺贝尔医学奖。

我们可以把红外线、可见光、紫外线都称作光。在这个波长频率范围内，$LC$ 回路已经无能为力了，只有原子外层电子跃迁才能产生。比如我们常见的日光灯管，就是汞原子受到电子撞击而激发，当它从激发态回到基态时，就会发出紫外线。紫外线再撞击荧光粉，就会使荧光粉发出可见光。

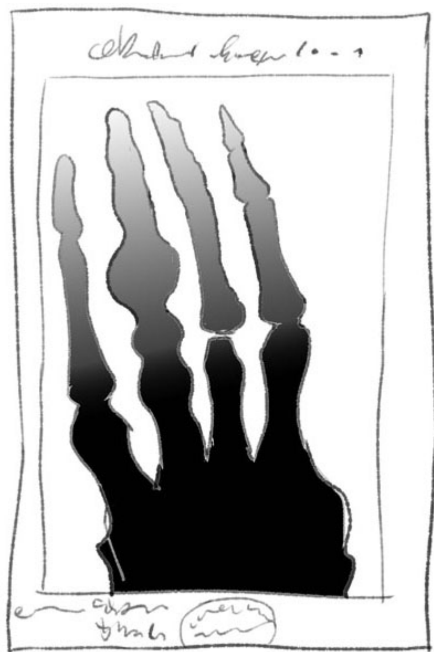
关于原子能级的理论，最早是玻尔为了解释氢光谱而提出的，他认为电子在原子外围存在不同的轨道，不同的轨道，其能量不同。电子在

每个轨道上运动时，都不会发出电磁波；只有电子在两个轨道之间跃迁时，才会吸收或者辐射电磁波。




当时人们还不能完全理解这个理论，因为这与传统电动力学的观点不同。不过现在人们已经认识到，在微观世界必须抛弃传统理论，而使用量子力学。

# 三、X射线



X射线最早是由伦琴发现的，所以也称为“伦琴射线”。伦琴不光发现了X射线，还拍摄了第一张X照片，是给他老婆的手拍摄的。在这张照片中，手的骨头和手指上的戒指都清晰可见。

X射线的波长比光还要短，能量更大。它是由于原子的内层电子跃迁产生的。目前X射线主要应用在安检和医疗上的透视。

# 四、伽马射线

伽马射线波长最短，频率最高，能量最大，穿透力最强。它是由于原子核跃迁产生的。就像原子的外层电子一样，原子核也具有能级。如果原子核从高能级回到基态，也会发出电磁波，因为这种能量非常巨大，所以形成的电磁波波长很短，于是，就形成了伽马射线。

伽马射线目前在医疗上用于杀死癌症细胞，在工业上主要用于金属探伤机。由于伽马射线能够使基因变异，很多电影都以此为主题展开想象，比如《曼哈顿博士》《神奇四侠》《绿巨人》，都是传说因为伽马射线照射而具有神奇的能力。

但事实上，原子弹爆炸时，杀伤力之一就是伽马射线。人在大剂量伽马射线照射下必死无疑。

综上所述，长波长、小频率的电磁波，是由于电荷周期性振动产生的。但光、X射线、伽马射线这种短波长、大频率的电磁波，则是由于原子或原子核进行能级跃迁时发出的，所满足的规律是量子力学。

# 2.11 量子是个什么玩意儿？

——量子力学的开创

十九世纪的最后一天，欧洲的物理学家齐聚一堂，迎接新世纪的来临。著名的科学家开尔文爵士惊叹于物理学的伟大成就，自豪地宣布物理学的研究已经走到尽头。




开尔文之所以这么说，是因为在那个时代，经典力学通过牛顿、拉格朗日、拉普拉斯等人的贡献，已经清楚地解释了物体之间的相互作用和天体运行规律，麦克斯韦电磁方程组将电与磁完美地统一起来，热力学统计物理可以解释分子的运动规律，仿佛物理学已经完全成熟了，没有什么重大的理论问题需要解决。以后的物理学家只需要将物理常数的精度提高几位就可以了。

但是，开尔文同时也说："在物理学晴朗的天空中，还飘着两朵令人不安的乌云。"他所说的这两朵乌云其一是指黑体辐射问题中实验结果与理论不符合，另一朵是指寻找光的参考系——以太的迈克尔逊-莫雷实验的失败。

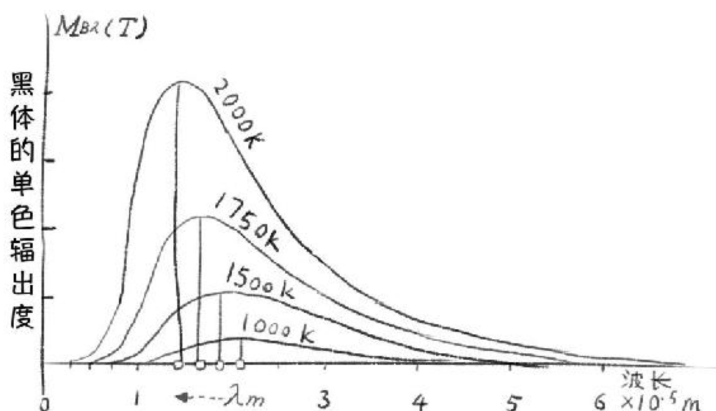
恰恰是这两朵乌云，发展成为二十世纪物理学最伟大的两个发现：

量子力学和相对论的诞生。人类认识到自己探索自然的道路还很漫长。

# 一、黑体

为了理解量子，我们首先介绍一下黑体。物理研究发现：一切物体都在吸收、反射和辐射电磁波。如果一个物体只吸收和辐射电磁波，不反射电磁波，这个物体就称为“黑体”。比如太阳就可以看作一个黑体，因为太阳的辐射特别强，辐射的电磁波强度远远大于反射的电磁波。

人们经过研究发现：黑体辐射的情况与物体的温度有关。




图中纵坐标是单位波长单位面积辐射功率，横坐标是波长。我们通过这个图可以发现两个结论：

第一，物体温度越高，辐射强度越大。黑体单位面积辐射能量与温度的四次方成正比，称为“斯特藩-玻尔兹曼定律”。人们根据这个规律计算了太阳表面温度大约是6000K。

第二，物体温度越高，辐射强度最大处的波长越短，称为“维恩位移定律”。比如炽热的铁块会发光，而且温度不同时，颜色也不同。有经验的铁匠可以根据铁块的颜色判断铁块的温度。

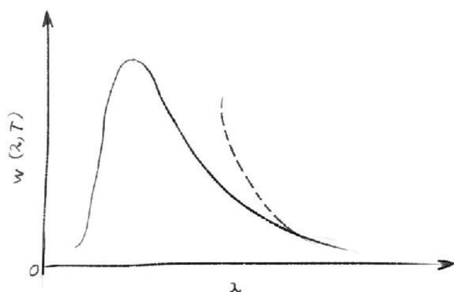
# 二、紫外灾难

但是，这两个定律都是实验规律，如何从理论上解释呢？

卡文迪许实验室主任瑞利从经典电动力学出发，推导出一个黑体辐射公式，即瑞利-金斯公式。

$$M_{B\lambda}(T)=\frac{2\pi c}{\lambda^4}kT$$ （瑞利-金斯公式）

不过，这个公式并不能符合实验结果。只有在波长比较大的时候，公式才与实验结果符合，在波长较小时，公式与实验结果偏差很大。




最可怕的是，当波长趋近于零时，瑞利公式的结果发散，辐射强度无穷大，这显然是很荒谬的。人们无法调和理论和实验结果，并把这个问题称为“紫外灾难”（这是因为紫外是比可见光波长更短的光，表示波长短时，实验结果与理论值不符）。

# 三、普朗克和量子



为了解释这个问题，许多物理学家提出了自己的见解。最成功的是德国科学家普朗克。左边是普朗克学习物理过程中相貌变化图。

普朗克在1900年提出：为了解释黑体辐射现象，必须做出一定的假设，这些假设可能与人们熟悉的物理学规律不同。

振动的带电粒子能量是一份一份的，每一份的能量都与振动频率有关，称为“一个能量子”，或简称为“量子”。

$$\varepsilon = h\nu \begin{cases} \varepsilon: \text{能量} \\ h = 6.63 \times 10^{-34}\,Js\ \text{普朗克常数} \\ \nu: \text{频率} \end{cases}$$

按照这个假设，普朗克推导出了黑体辐射的普朗克公式。

$$M_{B\lambda}(T) = \frac{2\pi hc^2}{\lambda^5} \cdot \frac{1}{e^{\frac{hc}{k\lambda T}} - 1}$$

这个公式与实验结果符合得非常好，黑体辐射问题得到完美解决。但是，许多物理学家并不能完全理解量子的概念，这与经典物理学的冲突使得科学界始终不能完全相信普朗克的假设。直到18年后，普朗克才获得诺贝尔奖。



但是，能量子的概念提出后，许多物理学家借用这个概念取得了丰硕的成果。例如1905年，爱因斯坦借用普朗克的观点解释了光电效应实验，并获得了诺贝尔奖。

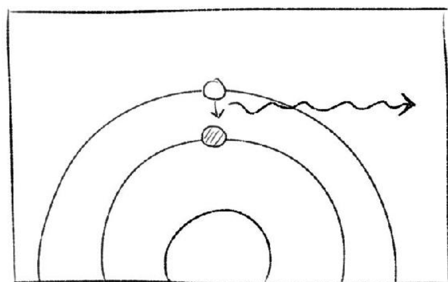
现在人们认识到：量子力学是统治微观领域的物理规律，它与宏观世界满足的规律不同。牛顿定律统治着宏观低速世界，量子力学主宰着原子量级的微观世界，而在高速时，我们又需要求助于相对论了。科学越发展，就越会发现更多未知的世界。

正如牛顿所说："我好像是一个在海边玩耍的孩子，不时为拾到比通常更光滑的石子或更美丽的贝壳而欢欣鼓舞，而展现在我面前的是完全未探明的真理的海洋。"

# 2.12 波还是粒子？这是个问题

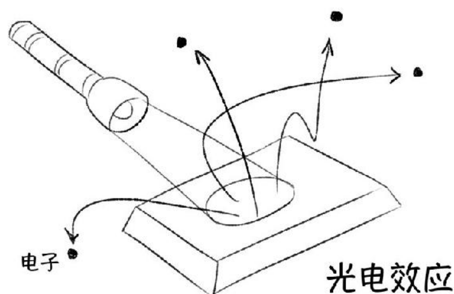
——波粒二象性

光到底是波，还是粒子？这在物理学界经历了长期的争论。牛顿是微粒说的代表人物，而惠更斯则认为光是机械波。经历了麦克斯韦、赫兹、托马斯·杨、菲涅耳等人的努力，人们逐渐认识到光是一种电磁波。十九世纪初，人们自以为已经认清了光的本性。



## 一、爱因斯坦：光的粒子性

十九世纪初，科学家赫兹发现了光电效应现象：紫外线照射可以使得锌板发射电子。




原本大家以为这是个平淡无奇的现象，因为光具有能量，可以将电子撞出。最初人们认为光的能量与光强有关，因此越强的光，越容易发生光电效应，但是这个想法却无法获得实验支持。实验发现光电效应是否发生与光的强弱无关，而与光的频率有关：频率越大，越容易发生光

电效应。




为了解释这个问题，爱因斯坦大胆借用了普朗克的观点。他认为，光的能量是一份份的，每一份称为“一个光量子”，或简称“光子”，光子的能量与频率的关系也满足普朗克公式。

比如，紫外线光子的能量就比可见光强，可见光的光子能量又比红外线强。因此，只有频率高的光，才能将电子撞出。光强并不表示每个光子的能量，而表示光子的个数。爱因斯坦通过这个关系完美解释了光电效应实验，并获得诺贝尔奖。

于是，在爱因斯坦提出了光子学说之后，人们认识到光不只具有波动性，也具有粒子性，于是就称为“波粒二象性”。爱因斯坦说：“好像有时我们必须用一套理论，有时候又必须用另一套理论来描述（这些粒子的行为），有时候又必须两者都用。”

# 二、德布罗意：粒子的波动性

既然电磁波是有粒子性的，那么粒子是否也有波动性呢？这个想法看似天方夜谭，一个苹果如何能跟波联系到一起？

但是自然界就是这么神奇，就好像法拉第发现了变化的磁场可以产生电场，麦克斯韦就联想到变化的电场也能产生磁场一样，一位年轻的法国学者大胆地预言：不只光具有波粒二象形，实物粒子也有波粒二象性。这就是法国学者路易·维克多·德布罗意。




德布罗意家族十七世纪以来一直为法国国王服务，1740年被授予公爵称号，由长子继承，后来家族中的每个成员又都获得神圣罗马帝国亲王称号。德布罗意的父亲是法国总理和外交部长，在德布罗意的兄长死后，他便成为了法国公爵和德国亲王。




德布罗意本来是学习中世纪欧洲历史的，并获得了文学学士学位。后来他碰巧阅读了一些有关物理的书，尤其是聆听了法国数学家庞加莱的报告后，突然对物理产生了浓厚的兴趣，转而攻读物理学博士学位。

读了几年书，德布罗意面临毕业问题。但是死活写不出论文来。他着急，他的导师朗之万更着急，因为面对这么个官二代谁也惹不起。可是写不出论文来肯定不能毕业，这可怎么办？

1921年，爱因斯坦借用普朗克的量子假说解释了光电效应实验，并获得了诺贝尔奖。爱因斯坦指出：光具有波粒二象性。

德布罗意灵机一动，在爱因斯坦的结论上添了一句话：不只是光，

所有的物质都具有波粒二象性。物质的粒子性由动量$P$ 代表（质量与速度的乘积），波动性由波长λ 代表，并且二者的乘积等于普朗克常数$h$。

$$\lambda = \frac{h}{P}$$

比如，一颗子弹的质量$m$ = 0.1kg，当它以$v$ = 300m/s的速度运动的时候，子弹的动量$P$ = $mv$ = 30kgm/s。这样子弹的波长$\lambda = \frac{h}{P} = \frac{6.63\times 10^{-34}\,\text{Js}}{30\text{kgm/s}} = 2.21\times 10^{-35}\,\text{m}$，这个波长如此之短，任何仪器都无法探测到，但它是存在的。

德布罗意写了一份论文交给了导师朗之万。

由于这份论文思想过于超前，朗之万也害怕过不了关，于是把论文写信寄给爱因斯坦，并特别提到：这是一个法国公爵面临能否毕业的重大政治问题，如果您能赞同他的观点，那么您以后来法国肯定会很受欢迎的。

爱因斯坦欣然明白了朗之万的意思，当即回信说论文的结论很有见解。于是在爱因斯坦的肯定下，德布罗意顺利毕业了。

过了几年，诺贝尔奖委员会开始审议当年诺贝尔奖候选人时，突然发现德布罗意说的好像是对的，于是决定授予他诺贝尔物理学奖。他是世界上第一个通过博士论文拿到诺贝尔奖的人。

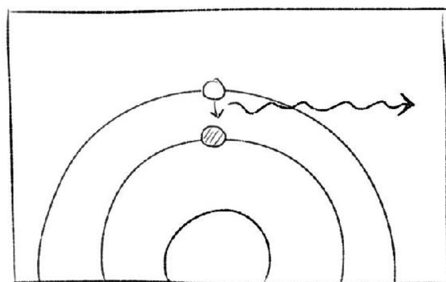


实际上，在此之前，量子力学教父级人物——丹麦物理学家尼尔斯·波尔在1913年提出了氢原子能量量子化模型。

波尔指出：电子在围绕氢原子运动时，轨道只能取某些特定的值。这些特定的值满足量子化条件：

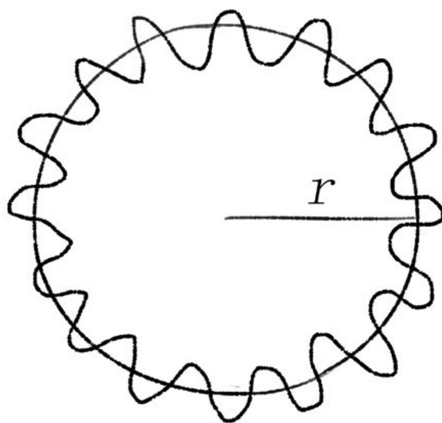
$$mrv = n\frac{h}{2\pi}$$

其中，$m$ 是电子质量，$r$ 是电子轨道半径，$v$ 是电子速度，$n$ 是一个整数，称为“量子数”，$h$ 是普朗克常数。



当电子在不同轨道之间跃迁时，氢原子就可以发射光子。波尔通过这个假设成功地解释了氢原子发光现象，并获得诺贝尔奖。

可是，为什么氢原子的运动要满足这个规律呢？1923年，德布罗意在《法国科学院学报》上连续发表了三篇论文，解释了波尔量子化的条件：电子轨道必须使得电子在原子核周围形成驻波。




所谓驻波，就是指电子的波长必须能够首尾相接。也就是说，电子

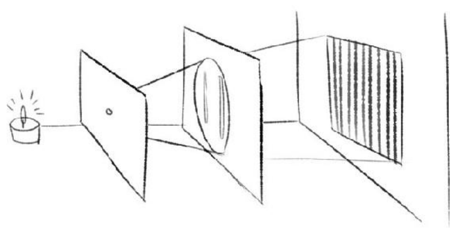
轨道周长必须是波长的整数倍。

$$2\pi r = n\lambda$$

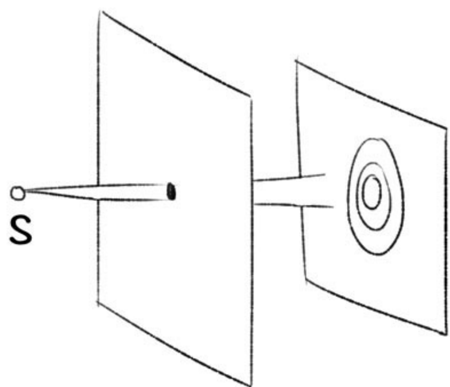
如果电子的波长与动量的关系满足$\lambda = \frac{h}{P}$，就可以得到$2\pi r = n\frac{h}{mv}$，就与波尔的结论保持一致了。

# 三、波粒二象性的实验证实

要证实一个物理理论，必须通过实验。既然粒子具有波动性，那么就应该能表现出波的特点，那就是干涉和衍射。干涉和衍射是指波通过障碍物时，传播方向会发生变化，造成障碍物后面出现不同于直线传播的图案。例如，双缝干涉就是光通过两个缝隙之后，在后面的屏幕上出现明暗相间的条纹。

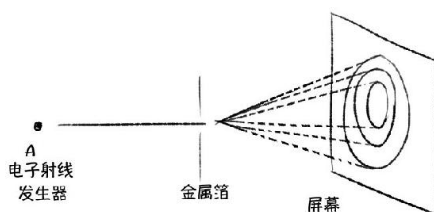


圆孔衍射就是光通过一个小孔之后在后面的屏幕上形成同心圆环。




干涉和衍射是波特有的，声波、水波、光波都具有干涉和衍射现象。那么要证实物质波的存在，就必须发现粒子的干涉和衍射现象。

终于，科学家G．P．汤姆孙成功地观测到了电子的圆孔衍射图样：将电子通过一个狭窄的缝，电子经过缝之后，居然表现出了光的特点——在屏幕上出现了衍射条纹。至此物质波的学说被人们证实了。




# 四、物质波的本质：概率波

那么，粒子的波动性本质到底是什么呢？

德国物理学家玻恩提出：粒子的波动性与经典的机械波不同，它并不表示振动形式的传递，而是表示粒子存在于各个不同位置的概率。也就是说，微观世界中的粒子并不一定在某个位置，而是存在一定的概率分布。在某些地方概率大，在某些地方概率小，用波函数就能描述这种差别。这种解释称为“哥本哈根诠释”。宏观物质——例如一个苹果，也具有波动性，只是由于波长太小，所以人们才会没有感觉。

德布罗意是个具有深刻洞察力的科学家，他从爱因斯坦的理论出发，通过大胆的猜想，成功地解释了波尔的理论，同时还预言了G．P．汤姆孙的电子衍射，并成功开拓了物质波理论，最终由波恩用概率理论解释，这一切都成为量子力学的重要内容。他宛如一座桥梁，连接了几个诺贝尔奖的成果。而相比于他所连接的几个成果，德布罗意的工作基本没有做过实验，理论计算也很少，而是凭借深刻的洞察力“轻而易举”拿到了诺贝尔奖。科学就是这样，在前人的基础上经过长期的思索和努力，使科学进步一点点。在这个过程中，有些人的工作异常辛苦，有些人却相对容易得多。

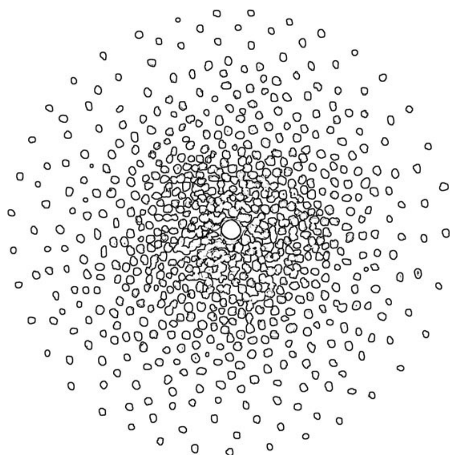
# 2.13　薛定谔的猫是死了还是活着？

——薛定谔的猫

这回要给大家讲一个科学史上最著名的思想实验——“薛定谔的猫”。这是一个把量子力学应用到宏观世界时出现的神奇悖论。



## 一、神奇的量子世界：叠加态和本征态



我们知道，量子的世界非常神奇，在量子的世界里，确定性被概率

取代了，无论是粒子的位置、能量，还是速度，都处于一种不确定的状态之中。

例如，氢原子是由原子核和核外的一个电子组成的，电子会围绕原子核高速运动。最初波尔在解释氢原子时，认为氢原子的电子存在不同的轨道。但是他发现，这种理论只对氢原子有效，稍微复杂一点的原子都无法解释。后来人们终于发现，电子并不存在确定的轨道，它的空间位置是随机的，于是人们画出了电子云，表示氢原子中的电子出现在各个不同位置的概率。

在德布罗意提出物质波的概念之后，波恩通过概率说解释了物质波和波函数的含义：波函数表示量子系统中某个事件的概率。

比如波函数$\Psi(r, t)$表示一个随着位置$r$和时间$t$演化的波函数，那么$|\Psi(r, t)|^2$就表示在位置$r$和时刻$t$找到粒子的概率。波尔等人认为这种观点是正确的，人们把这种对于波函数和量子力学基本问题的解释称为“哥本哈根诠释”。

更为神奇的是，因为量子系统的概率诠释，我们在没有进行观测时，不能确定一个粒子的位置和速度等信息，因此量子系统就处于一种叠加态。例如粒子既可能在$A$处，也可能在$B$处，它就处于$A$和$B$两处的叠加态；一个原子核可能衰变，也可能没有衰变，它就处于衰变和未衰变的叠加态。

我们要知道这个粒子到底处于$A$，还是处于$B$，或者要知道一个原子核到底衰变了没有，就需要进行观测。我们可能发现粒子在$A$处，也可能发现粒子在$B$处，我们称为“叠加态坍塌成了本征态”。似乎我们的观测是会影响结果的，因为在观测之前，粒子究竟在哪里，是不确定的，而观测之后，粒子立刻选择了$A$位置或$B$位置，这个过程就是在我们观测的一瞬间发生的。而且从此之后，粒子的状态就确定了。

这种观点与宏观世界是相违背的。比如我们到酒吧里去喝酒，服务员为我们倒了一杯酒，可能是啤酒，也可能是葡萄酒，我们可以通过品尝知道这杯酒是什么。然而，就算我们不去品尝，我们也知道这杯酒只可能是啤酒或者葡萄酒，虽然我们不知道这杯酒是什么，但是服务员知

道。就算服务员也不知道，我们何时品尝，或者谁来品尝，这杯酒的测量结果都不会有不同。因此宏观世界不是处于叠加态，而是处于一个本征态的。




但是量子世界里的酒却不是这样。这杯酒处于啤酒和葡萄酒的叠加态，既有可能是啤酒，也有可能是葡萄酒。现在世界上任何一个人都不知道这杯酒是什么，而且在不同的时刻，不同的人来品尝这杯酒，结果都可能不同。

由于量子力学中有太多与常识相违背的结论，所以许多人对量子力学产生了怀疑。许多人认为量子力学是一个不完备的理论，它只是一个更深刻的物理结论的某一个侧面。这其中就包含着量子力学的许多创立者，如爱因斯坦说“上帝不掷骰子”，薛定谔也提出了“薛定谔的猫”。

# 二、薛定谔的猫

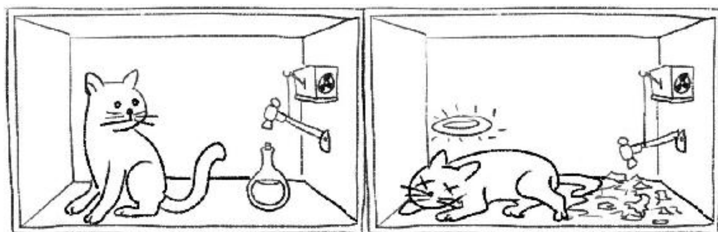


奥地利物理学家埃尔温·薛定谔是量子力学的奠基人之一。他在1926年提出了薛定谔方程，用以描述量子态的波函数随着时间的演化，并获得诺贝尔奖。

但是，他却认为量子力学并不是一个完备的理论。尤其是在宏观世界中会有许多与量子力学相违背的事实。他为了把这个事实描述得更加清晰，就提出了薛定谔的猫这个最让物理学家们头痛的思想实验。

把一只猫关在一个封闭的铁容器里面，并放入少量的放射性物质。这些物质在1小时内有50%的概率发生衰变，还有50%的概率不发生衰变。如果物质发生了衰变，旁边的探测器会探测到，这样会通过一个装置启动一个榔头。榔头会打碎一个装有有毒物质的瓶子。瓶子里的有毒物质扩散到容器里，就会把猫毒死。如果物质没有衰变，那么有毒物质不会扩散，猫就活得好好的。

由于量子系统处于叠加态，因此在人们没有打开盒子看的时候，这些放射性物质处于衰变和没有衰变的叠加态之中，这就使得这只猫处于一种既活又死的叠加态之中。只有打开盒子进行观测，在这一瞬间，叠加态会瞬间坍塌成本征态，这只猫就从一个既死又活的状态立刻变为活的或者死的猫。






有人可能会认为，不打开盒子也没关系，我们可以在盒子上安装玻璃看着这只猫啊。要知道，任何的观测行为都会影响实验。比如安装玻璃我们能够看到内部，这是因为有光射入了盒子再反射出来，这些光子就会影响量子系统，所以不能完成实验。猫要处于真正的叠加态之中，必须排除外界的任何干扰，因此人们无法观测。

这只猫的出现让物理学家们抓狂了。人们差一点就相信了量子力学和哥本哈根诠释，但这个美好的愿望被一只猫打击得粉碎。

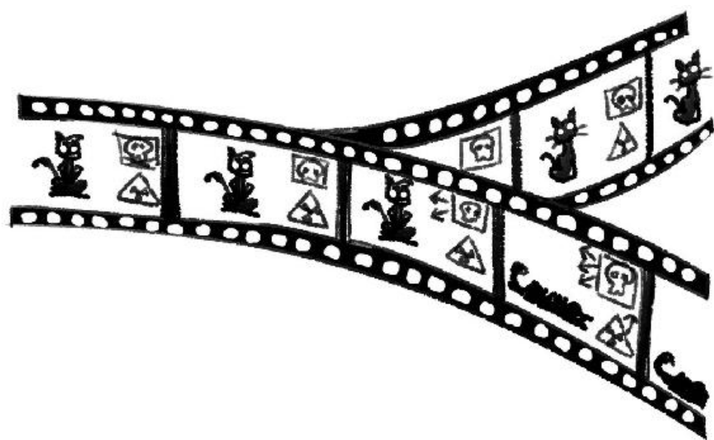
薛定谔通过这个实验向世界阐述：量子力学只是某个更深刻物理原理的侧面。

# 三、平行世界

那么，“薛定谔的猫”究竟该如何解释呢？现在的科学界对此还没有统一的观点。有些科学家甚至提出了许多神奇的理论来解释它。

1957年，科学家休·艾弗雷特提出了著名的“多世界诠释”。他认为，在进行薛定谔的猫的实验时，箱子里原本就有两个世界。这两个世界在箱子外的情况完全相同，只是一个世界里箱子里有只死猫，而另一个世

界里箱子里有只活猫，只不过这两个世界是纠缠在一起的。当我们打开箱子进行观测时，这两个世界就会发生分离，从此之后，各自变为一个新的世界，而且彼此毫无影响。



这种想法多么的神奇！科幻作者们为此欣喜若狂，一大批科幻小说和电影都通过这个理论演绎平行世界的精彩故事，比如《彗星来的那一夜》就是部不错的电影。可是这种想法又如此不可思议，让大部分科学家不愿承认它。薛定谔的猫至今仍是一个无法完全理解的“怪物”。

量子力学究竟是不是完备理论？人们对它还存在很多争论。著名的物理学家、诺贝尔奖获得者、最幽默的物理学家费曼曾经说过：“我想我可以有把握地讲，没有人懂量子力学！”

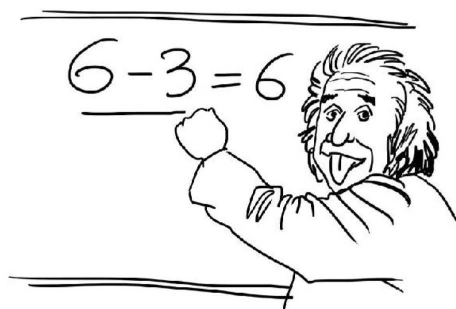
# 2.14　黑洞是黑色的吗？

——从爱因斯坦到霍金

剑桥大学卢卡斯数学教授史蒂芬·霍金在2018年3月14日去世了，这个日子正是爱因斯坦诞辰的日子。霍金21岁时因为患病而全身瘫痪，不能言语，手部只有三根手指可以活动。但是在这样的情况下，他凭借惊人的毅力提出了广义相对论的奇性定理和黑洞面积定理，提出了黑洞蒸发理论和无边界的霍金宇宙模型，同时也凭借科普著作《时间简史》《果壳中的宇宙》等成为世界科普大师，去世后，霍金被安葬在伦敦西敏寺·牛顿的墓旁。

霍金研究一生的黑洞到底是什么呢？为了了解这个问题，首先我们必须说回到爱因斯坦。

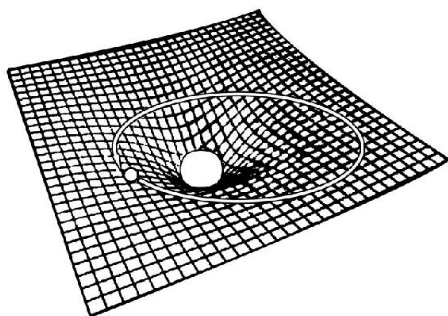
## 一、爱因斯坦和广义相对论




爱因斯坦的主要贡献是解释了光电效应、提出了狭义相对论和广义相对论。这些贡献中，又以广义相对论最为精彩，他是继牛顿和麦克斯韦之后人类认识世界的第三次飞跃，是我们研究宇宙的有力工具。现代宇宙学的基本观点：大爆炸理论、宇宙膨胀、黑洞、引力波等问题，无一不是通过爱因斯坦广义相对论进行解释的。

在广义相对论中，爱因斯坦把牛顿的“质量引起引力重力场”的观点

引申为质量引起时空弯曲，并通过时空弯曲求解物体的运动规律。例如，牛顿认为，地球围绕太阳运动是因为太阳对地球有万有引力，但是广义相对论却把这个过程解释为太阳的质量很大，引起了周围空间的弯曲。地球的圆周运动在这个弯曲空间中实际上是一条“直线”。



在这个理论里面，爱因斯坦提出了引力场方程：

$$G_{\mu\nu} = R_{\mu\nu} - \frac{1}{2} g_{\mu\nu} R = \frac{8\pi G}{c^4} T_{\mu\nu}$$

这个方程看起来不复杂，但它是个张量方程的简化写法，整个展开非常麻烦。通过求解这个方程，一些科学家得到了一些很惊艳的结果。

例如：德国天文学家史瓦西通过计算，得到了引力场方程的一个特殊解。

他发现，这个解很奇异，有很多奇妙的性质。于是他就把这个解叫作黑洞，并指出宇宙中可能存在这种比较奇怪的天体。

宇宙中究竟有没有黑洞呢？人们开始了一个世纪的寻找，直到今天，人们依然把黑洞的存在定性为“可能”。

# 二、霍金的贡献



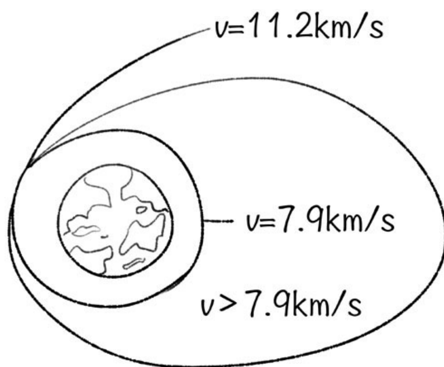

那么霍金又做了什么呢？

霍金对这个解又进行了一些尝试，在爱因斯坦广义相对论的基础之上，又引入了量子场论，并通过量子场论的分析，对黑洞的性质进行了更加详细的描述。比如霍金提出了奇点理论、黑洞蒸发理论、证明了黑洞面积定理，等等。霍金的数学功底非常强，他的理论都是通过完美的数学表达式得到的结果。但到目前为止，以上的理论还只是推测。黑洞是否真的存在，还是一个谜。霍金也没有因此获得诺贝尔奖。

# 三、奇妙的黑洞

黑洞到底有哪些神奇的性质呢？

首先，黑洞质量非常大，以至于连跑的最快的物体——光——都无法逃逸。为了理解“逃逸”这个概念，我们从比较经典的例子出发，讨论一下宇宙速度。




牛顿曾经说：一个炮弹在地面发射，如果速度比较慢，就会落回到

地面。

如果速度达到一定值，就不会落地，而是环绕地球运动。这个速度称为“第一宇宙速度”，它的大小是7.9km/s。

如果这个速度继续增大到11.2km/s，物体就再也不会回到地球，而是逃逸到无穷远处，这个速度也就被称之为“逃逸速度”。

在经典物理里，逃逸速度与星球质量和半径有关：

$$v_2=\sqrt{\frac{2GM}{R}}$$

其中，$G$ 是万有引力常数，$M$ 是星球的质量，$R$ 是星球的半径。

我们会发现：如果星球的质量$M$ 越来越大，而星球的半径$R$ 越来越小，逃逸速度$v$ 就会增大。如果这个速度增大到光速$c$ ，那么任何物体都无法逃脱星球的引力。

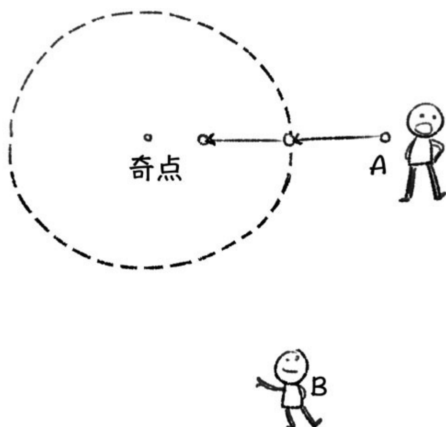
于是，我们利用公式$c=\sqrt{\frac{2GM}{R}}$ 得到$R=\frac{2GM}{c^2}$ 。

这个半径大小就称为“史瓦西半径”。也就是说，如果一个星球的半径小于史瓦西半径，就连光也逃不出去。

其实，以上只是一个类比过程，史瓦西半径是从广义相对论那个非常复杂的方程得到的，但是恰巧结果与经典物理的表达式完全一样，为了便于大家理解，我们才姑且这样处理。如果我们把地球的质量$M = 6\times10^{24}$ kg代入，就会得到地球的史瓦西半径$R = 0.01$m。



黑洞最引人入胜的地方是它会造成强烈的时空弯曲。时空弯曲的概念不太好理解，我们从一个简单的例子来说明：假如由一个引力很大的天体构成了黑洞，那么它的外界有一个圆圈：视界，其半径大小就是史瓦西半径。在视界之外，物体还有机会逃逸黑洞，但是一旦进入视界，即使以光速运动的物体，也无法逃脱黑洞。




假设有一个人$A$ 在匀速接近黑洞，$A$ 会认为自己在匀速直线运动。但如果在遥远处有一个人$B$ 观察，他的感觉却与$A$ 不同：在$B$ 看来，$A$ 的速度越来越慢。最终当$A$ 接近视界边缘的时候，$B$ 就会发现$A$ 静止不动了。在$B$ 看来，$A$ 永远都不会进入视界，而是会成为一座在视界边缘的丰碑。这就是引力造成的时间变慢效应：距离黑洞越近，时间越慢。

不仅仅是黑洞，大质量物体都会造成周围的时间变慢，大家应该看过一部科幻电影《星际穿越》的一个场景：几个宇航员到一个星球仅仅停留了几小时，而在外围空间轨道等待他们的宇航员却在飞船上待了20多年。




回到黑洞问题，在$A$ 自己看来，他是可以进入视界的。而且当$A$ 进入视界之后，就再也出不来了，因为即便他以光速运动，也无法从视界逃逸。视界内的所有现象都和视界外无关，视界内的任何信息都无法发送到视界外，所以我们可以说视界内和视界外是两个世界。

更为有趣的是，在黑洞外，时间是单向的，空间是双向的。也就是说，人不能回到过去，但是可以往前运动，也可以往后运动。不过，在黑洞里面，由于时空扭曲得特别厉害，时间会变成双向，可以回到过去。但空间却变成单向的了，不能往回走，物体只能向前，向着黑洞中的一个点——奇点运动。

奇点是黑洞中心一个密度无穷大的点，在这个点上，一切物理规律和数学规律都失效了，类似于用1除以0这个表达式。所以我们才称之为“奇点”。

$$Y=\frac{1}{\quad}$$

在物体靠近奇点的过程中，距离奇点越近，引力差越大，最终物体

会由于两端受力差太大而被撕裂。

黑洞如此奇妙，吸引了无数科学家的目光。但是黑洞不会发出任何光，我们怎么寻找它呢？霍金预言：黑洞视界附近可以产生正负物质对，正物质放出形成霍金辐射，负物质被吸进黑洞造成黑洞质量损失，这个过程称为“黑洞蒸发”。所以根据霍金的理论，人们认为黑洞也存在寿命，而且可以通过观察霍金辐射来寻找黑洞。

我们如果想销毁一份文件，可以用碎纸机把它剪碎，再用火烧掉纸屑，但即便这样，理论上说，信息并没有消失，如果我们使用巧妙的方法进行还原，还是可以读到原来的信息。但是黑洞不是这样，它吞噬的所有信息都会永远消失，虽然黑洞在蒸发，但是黑洞的蒸发是不携带任何信息的。黑洞是宇宙的超级碎纸机。黑洞理论依然是宇宙学中最神奇的内容，期待着更多科学家进行讨论。

# 2.15　如何制作原子弹

——质能方程与核反应

核武器是人类现在掌握的威力最大的武器，1945年，美国在二战最后阶段向日本的广岛和长崎投放了两颗原子弹，耀眼的闪光和巨大的轰鸣之后，20余万人死伤，城市被夷为平地。

从此，人们认识到核武器的巨大威力。二战之后，以美、苏两国为首的冷战集团疯狂进行军备竞赛，原子弹数量和爆炸力直线提升。美国在日本投放的“小男孩”和“胖子”两颗原子弹不过是2万吨TNT当量，可是苏联爆炸的沙皇核弹却有5000万吨TNT当量。爆炸后的蕈状云宽达将近40公里，高达60公里，比珠穆朗玛峰还高7倍多；爆炸产生的热辐射甚至可以让远在170公里以外的人受到3级灼伤，爆炸的闪光还能造成220公里以外的人眼睛剧痛，甚至灼伤。




核武器为什么有这么大威力？它又是如何制作的呢？

## 一、质能方程

为了解释这个问题，我们首先要从1905年——物理学上第二个奇异年——说起。

之所以说是第二个，是因为第一个物理学奇异年是1666年。在那一年，牛顿回到乡下老家躲避瘟疫，结果在很短的时间内，他就发明了微积分，发现了光的色散，并提出了万有引力，奠定了近代数学、光学和力学的基础。人们认为这是前无古人后无来者的事，就把1666年称为“牛顿奇异年”。

然而，1905年，还是专利局小职员的爱因斯坦连续发表了六篇论文，阐述了分子动理论、相对论、光电效应等问题，这六篇论文，每一篇都在科学史上具有举足轻重的地位。于是，后人将这一年称为“爱因斯坦奇异年”。

在其中的一篇论文《物体的惯性同它所含的能量有关吗》中，爱因斯坦阐述了他的观点：物体的能量与它的质量有关。这就是著名的质能方程：

$$E = mc^2$$

其中，$E$ 表示能量，$m$ 是物体的质量，而$c$ 表示真空中光速$c = 3\times10^8$ m/s。

如果一个物体的质量$m = 1$kg，按照这个方程计算，这个物体的能量有$9\times10^{16}$ J，大约相当于2000吨TNT炸药爆炸释放的能量。

但是，怎么从没见过1kg的物体有这么大威力呢？这是因为只有在核反应过程中，这个公式才会用到。

# 二、链式反应

在二十世纪初，人们已经认识到原子是由原子核和核外电子构成的。发现这种结构的科学家卢瑟福特别喜欢使用α粒子（氦原子核）轰击各种其他物质。

例如，在1917年，他使用α粒子轰击氮（N）原子核，产生了氧（O）原子核和质子（p）。

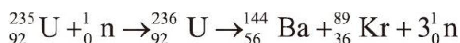


$$_{2}^{4}He+_{7}^{14}N\rightarrow_{8}^{17}O+_{1}^{1}p$$

这是人类第一次观测到原子核的变化。由于这是人为使用一个粒子撞击原子核促使原子核发生变化，因此称为“人工核反应”。

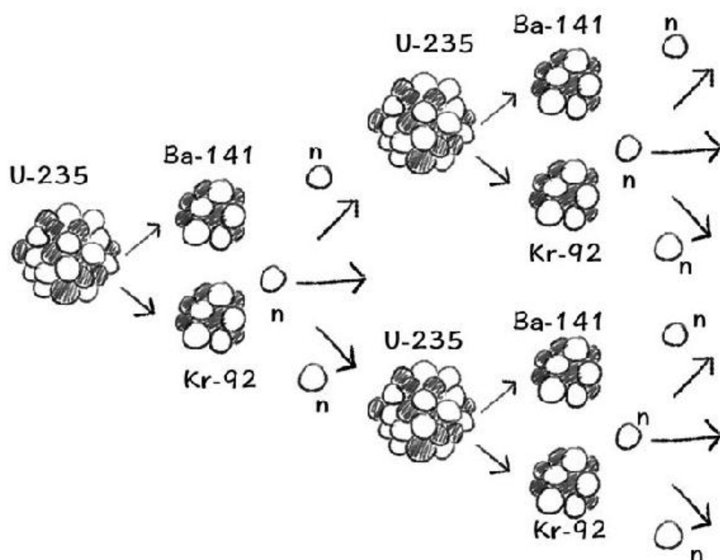
1938年，德国科学家奥托哈恩第一个发现了铀的裂变反应。

他使用中子（n）轰击铀（U）原子核，发现了钡（Ba）元素，反应方程是：



$$_{92}^{235}U+_{0}^{1}n\rightarrow_{92}^{236}U\rightarrow_{56}^{144}Ba+_{36}^{89}Kr+3_{0}^{1}n$$

在这个反应中，一个中子射入撞击到铀原子核上，使铀235转化为铀236，又继而变为钡144、氪89和3个中子。如果铀的体积和质量足够大，那么这3个中子会继续和3个铀235进行核反应，形成9个中子……如此这般，就会形成一种“雪崩效应”，短时间内产生许多次核反应。




人们把这种反应称为“链式反应”。在这个反应过程中，质子和中子的总量并不发生变化，但是平均每个质子和中子的质量却减少了$\Delta m = 0.3578\times10^{-27}$ kg，也就是发生了质量亏损。根据爱因斯坦的质能方程，这一部分质量亏损会转化成相应的能量$\Delta E = \Delta mc^2 = 201\text{MeV}$。

当时，世界正笼罩在第二次世界大战的阴云之中。纳粹组织了以著

名科学家海森堡为首的科学家团队，进行原子弹的研发。可是，进展却十分缓慢，原因之一是一大批有犹太血统的科学家，如爱因斯坦、赫兹等人，都被希特勒吓跑了，纳粹的内耗也相当严重，无法集中全部资源进行统一的研究工作。

# 三、曼哈顿计划

在日本偷袭美国珍珠港之后，美国被正式卷入战争。美国已经是当时世界第一的工业强国，工业实力有目共睹，珍珠港事件又使得美国全国上下同仇敌忾。在得知纳粹德国正在进行原子弹研制计划时，许多美国军方的科学家都建议抢在希特勒之前研制成实用的原子弹。为此，他们怂恿爱因斯坦帮忙。

爱因斯坦是美籍德国犹太人，当时已经50多岁了，在世界上享有巨大的声誉，纳粹上台之后，他逃到了美国。他写信给罗斯福总统说：如果纳粹首先研制出原子弹，那对全世界都是一个灾难。在爱因斯坦的号召下，美国政府集合了以奥本海默为首的一大批科学家，开展了历时3年的“曼哈顿计划”。

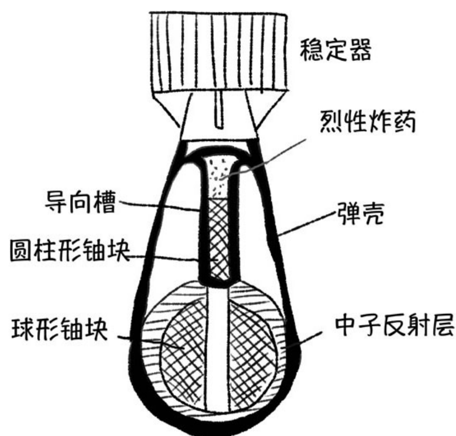



在原子弹制造的过程中，最难的问题是如何将铀进行提纯。原子弹的原料铀235在自然界的铀中只占0.7%，其他的是它的同位素铀238。为了使链式反应能够发生，铀235的浓度必须达到某个值以上。因为如果浓度太低，裂变放出的中子无法撞击到下一个铀235原子核上，链式

反应无法继续进行。“曼哈顿计划”大部分时间都在研究如何进行铀浓缩，到二战结束，美国提纯的铀刚刚够制造一颗实验炸弹和两颗实战用弹。

不仅仅是铀的浓度要够高，铀的体积也需要够大，称为“临界体积”。因为原子核相比于原子非常小，如果铀的体积不够，产生的中子很可能在原子核外的空间飞出去，链式反应也无法继续进行。

一个简单的原子弹模型如下：




在弹壳中有两块浓缩铀235，二者之间有一块柱状铀棒，三者都没有达到临界体积，因此不会爆炸。在原子弹引爆时，柱状铀块上方的炸药爆炸，将铀棒推入两块铀之间，三者合在一起之后体积超过临界体积，原子弹爆炸。在铀块外面的中子反射层用来反射中子，以减小临界体积。

原子弹的巨大威力造成了无差别攻击，使日本军人和普通民众同样承受了巨大伤亡。爱因斯坦获知这一消息之后后悔不已，说自己把原子弹从一个疯子（纳粹）手中夺下来，又交给了另一个疯子（美国）。

# 四、热核聚变

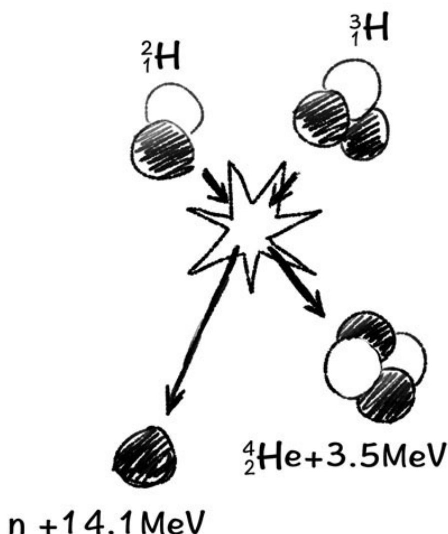
不仅如此，二战之后，整个世界笼罩在核阴影之中。美、苏等国家不仅疯狂制造原子弹，更制造出一种威力更大的核武器——氢弹。

人们发现：氢元素的两种同位素氚（${}^{2}_{1}\mathrm{H}$）和氚（${}^{3}_{1}\mathrm{H}$）原子核距离非常近时，会变为1个氦（He）核和1个中子（$n$），核反应方程是：



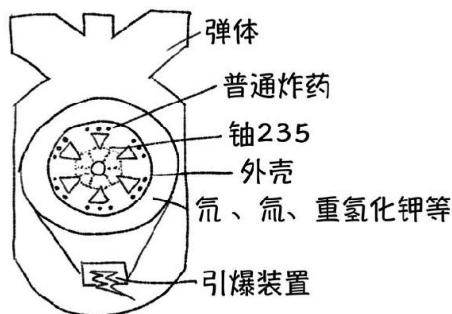
$$ {}^{2}_{1}\mathrm{H} + {}^{3}_{1}\mathrm{H} \rightarrow {}^{4}_{2}\mathrm{He} + {}^{1}_{0}\mathrm{n} $$

在这个过程中，原子核质量会减少$\Delta m = 0.0189\mathrm{u}$，同时释放能量$\Delta E = 17.6\mathrm{MeV}$。与同质量的铀裂变相比，聚变放出的能量要大很多。太阳发光的核反应原理就是这样。




但是，要实现聚变，首先要将氚核和氚核之间的距离减小到$10^{-15}\ m$量级，由于原子核之间带正电，这个过程需要巨大的能量。人们想到使用加热到几百万开尔文的办法引发聚变，因为在极高温度时，原子动能非常大，原子核可以凭借巨大的动能克服库伦排斥力。

用什么办法才能获得这么高的温度呢？普通炸药是无法达到这个温度的。于是人们想到了原子弹。




一个简单的氢弹原理是这样的：在聚变材料内部，有普通炸药和铀235。引爆氢弹时，首先将普通炸药引爆，造成铀235聚集在一起，使之超过临界体积，原子弹爆炸。裂变反应释放的巨大能量产生高温引发聚变，聚变开始之后，自身产生的热量就可以维持反应一直进行下去。氢弹爆炸时中心温度可达100亿度，比太阳的温度还要高。



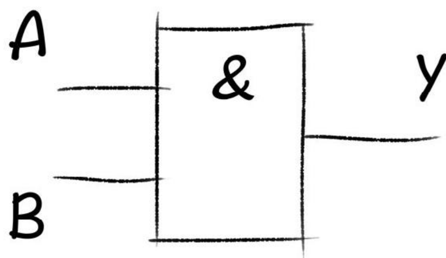

冷战时，美、苏两国储备的核武器可以把地球毁灭几十次，人们说：如果核战争爆发，那么下一次人类的战争可能只能使用木棒了。可也正是因为核武器有“同归于尽”的巨大威力，没有人敢轻易使用核武器，所以在日本遭受原子弹袭击之后，70多年来，核弹都没有再次应用于战场。

同时，人们利用核裂变技术制造了用于发电的核电站，相比于普通的火电站，核电站燃料消耗小，对环境的污染小，而且几乎是一种“取之不尽”的能源。

# 2.16　芯片为啥这么难做？

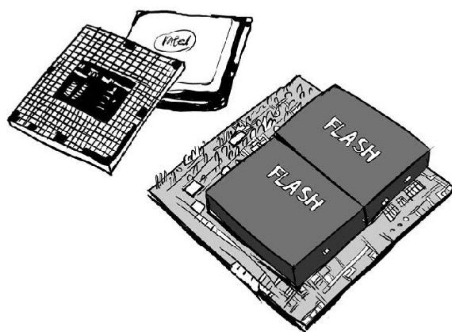
——芯片的基本原理与制作

我们都知道，小到手机，大到航天飞机，现在生活中方方面面的电子设备都需要一种叫作芯片（chip）的小玩意儿。芯片到底是什么呢？




## 一、芯片是什么？

简单来说，芯片是一个小型化的集成电路，在很小的体积上制作了许多的半导体元件，并且利用这些半导体元件实现一定的功能。比如，电脑中的CPU就是芯片，可以进行运算处理；电脑中的闪存FLASH也是芯片，可以实现信息的快速存储。




芯片的作用巨大，研发芯片需要许多高级人才，并且需要长时间、高投入，在之前，由于我国还属于发展中国家，芯片产业并没有进行大量的投入，造成现在我国芯片国有化率特别低。2018年4月16日，美国政府宣布对中国某企业实行禁售，一下子卡住了我们的脖子。

## 当前中国核心集成电路的国产芯片占有率

| 系统 | 设备 | 核心集成电路 | 国产芯片占有率 |
|---|---|---|---|
| 计算机系统 | 服务器 | MPU | 0% |
| | 个人电脑 | MPU | 0% |
| | 工业应用 | MCU | 2% |
| 通用电子系统 | 可编程逻辑设备 | FPGA/EPLD | 0% |
| | 数字信号处理设备 | DSP | 0% |
| 信息装备 | 移动通信终端 | Applicaton Processor | 18% |
| | | Communication Processor | 22% |
| | | Embedded MPU | 0% |
| | | Embedded DSP | 0% |
| | 核心网络设备 | NPU | 15% |
| 内存设备 | 半导体储存器 | DRAM | 0% |
| | | NAND FLASH | 0% |
| | | NORFLASH | 5% |
| | | Image Processor | 5% |
| 显示及视频系统 | 高清电视/智能电视 | Display Processor | 5% |
| | | Display Driver | 0% |

中国并不是没有过大力发展芯片的时候。20年前，中国就提出要大力发展芯片产业。在这个大环境下，2003年，上海交大软件与微电子学院院长陈进制作了一款芯片：汉芯一号，经过测试达到了世界一流水平，陈进也因此获得了无数荣誉。

2006年，陈进的一位研究生在清华大学水木BBS爆料：所谓的“汉芯一号”，其实是陈进在美国购买摩托罗拉飞思卡尔56800的芯片后，雇佣民工将表面的字样用砂纸磨掉，再找浦东的一家公司打上“汉芯一号”字样，并加上Logo，以此骗取政府一亿一千万元人民币的科研经费。

经过调查属实，陈进被取消了一切荣誉，并被上海交大开除。但是这件事却成为了中国芯片产业的分水岭，从那之后，芯片项目的审查非常严格，十几年里，中国的芯片产业并没有获得预期的进步。

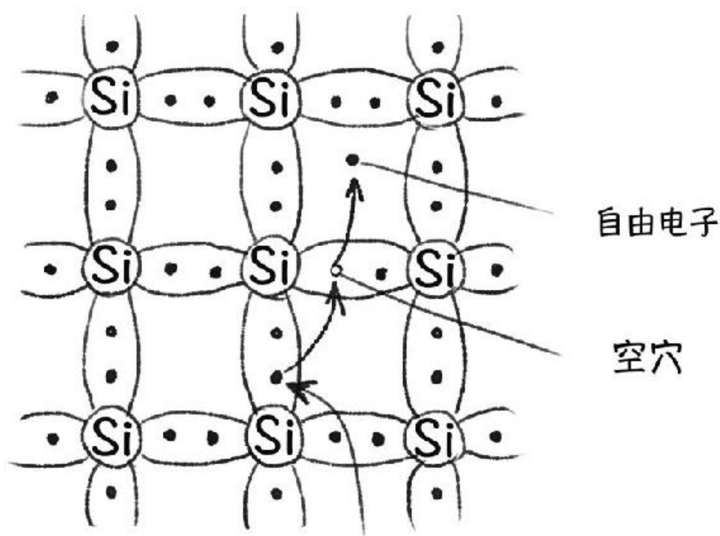



# 二、二极管

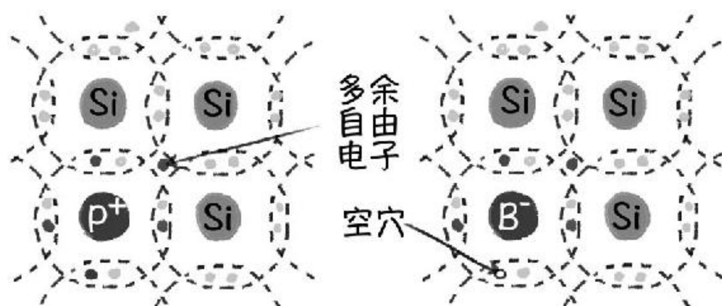
为了带大家了解芯片的基本原理，我们先从一个比较简单的半导体元件二极管说起。

二极管是一种基本的半导体元件，它的基本结构是PN结。

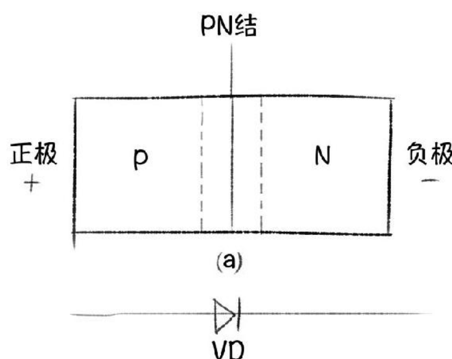
我们知道，硅是一种半导体材料。它的外围有四个电子。正常情况下，硅是不带电的。但是如果某一个电子飞出原子的束缚成为自由电子，原来电子的位置就形成了一个“空穴”，空穴是带正电的。



与硅不同的是，硼元素周围有三个电子，相比于硅少一个电子。如果把硼元素掺杂到硅元素中，硅与硼之间就会形成一个电子空位，相当于一个正电荷。这种半导体就称为“空穴导电型半导体”，或称“P型半导体”。相反，磷这种元素，最外层有五个电子，如果把磷掺杂到硅中，那么磷与硅之间就会多出一个电子，这种半导体就称为“电子导电型半导体”，或称“N型半导体”。




那么，假如我们在一块半导体硅上掺入三价的硼和五价的磷，各占一半，这样就形成了一种基本的结构：PN结。

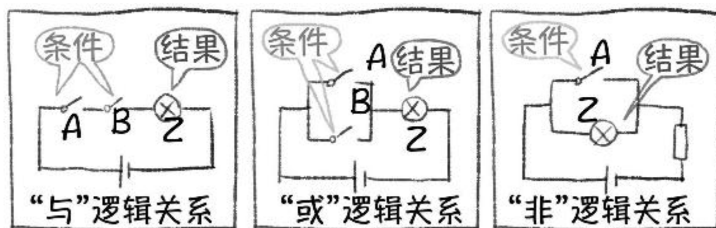



PN结中央会出现相互渗透的耗尽层，由于一些物理机制，造成PN结只允许电流从P端向N端流动，却不允许N端电流向P端移动。这就是二极管的单向导电机理。

# 三、与门、或门和非门

有了二极管，我们就可以实现一定的逻辑运算。基本的逻辑运算有“与”“或”“非”三种。为了理解这三种关系，我们用一个简单电路做个

比方。

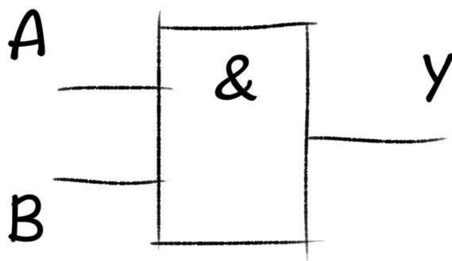



比如，两个开关与灯泡串联，只有两个开关都闭合时，灯泡才会发光，这就是“与”逻辑。如果两个开关并联，再与灯泡串联，只需要闭合一个开关，灯泡就会亮，这种关系就是逻辑“或”。如果开关与灯泡并联，开关闭合时灯泡被短路，灯泡不发光，反而是开关断开时，电流会流过灯泡，灯泡发光，就称为逻辑“非”。

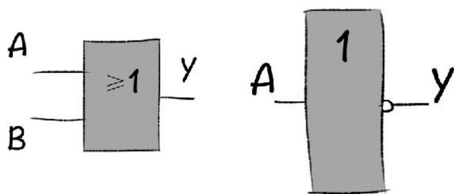
在逻辑电路中，用高电压和低电压分别表示两种不同的状态，与二进制中的“1”和“0”对应。这样一来，逻辑“与”的关系就是：如果$A$ 和$B$ 两个输入端都是高电平，则输出$Y$ 是高电平；只要有一个输入端是低电平，输出就是低电平。它的真值表和符号如下：

| A | B | A与B |
|---|---|---|
| 1 | 0 | 0 |
| 0 | 1 | 0 |
| 0 | 0 | 0 |
| 1 | 1 | 1 |

我们把能够实现这种逻辑关系的电路称为“与门”，符号如下：



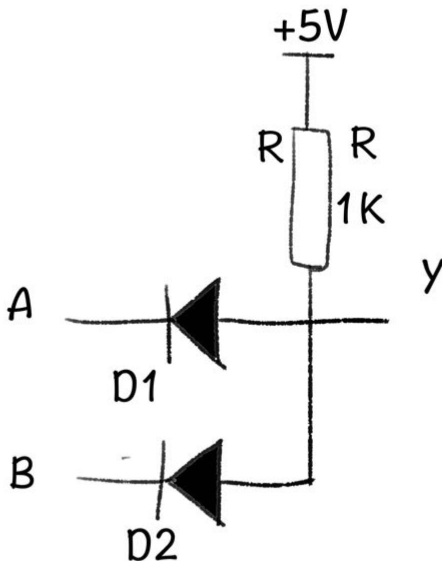

同样，我们还有或门和非门，符号如下：




同样地，逻辑“或”和逻辑“非”也有类似的关系。

# 四、门电路的物理实现

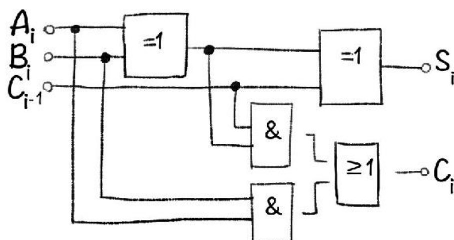
那么，与门如何进行物理实现呢？这就要用到二极管了。我们用接近5V的电压表示1，用接近0V的电压表示0。通过两个二极管按照下面的方法进行连接。




我们会发现：如果$A$ 和$B$ 都输入5V（高电平），那么整个电路各部分电压都是5V，于是$Y$ 端也是5V（高电平）。但是如果$A$ 输入0V（低电平），这样二极管$A$ 就会导通，由于二极管正向导通电压很小，这样$Y$ 就接近0V（低电平）。同样，$B$ 端输入0V时，$Y$ 端也输出低电平。这样就实现了逻辑与运算。或门和非门也有类似的结构。

利用“与门”“或门”和“非门”，人们就可以实现各种各样的计算需要。比如，我们可以利用它们实现一位数的加法，它的门电路逻辑图如

下：




计算机需要实现的功能比一位数加法要复杂得多，所以需要的门电路也复杂得多。第一代计算机使用更为原始的电子管组成，世界上第一台通用计算机“ENIAC”于1946年在美国宾夕法尼亚大学诞生，它是一个庞然大物，用了18000个电子管，占地170平方米，重达30吨，耗电功率约150千瓦，每秒钟可进行5000次运算。而且每几分钟都有电子管损坏需要更换，编程也非常复杂。

而现代的CPU把几十亿个晶体管放在一个指甲盖大小的晶片上，每秒可以实现几十亿次运算。显然，每一个晶体管都非常的小。这么小的结构是如何制作的呢？

# 五、芯片的制作流程

芯片的制作可以分为三个阶段。

第一个阶段是设计。就是研究为了实现某些功能，芯片中的各种半导体元件应该如何进行排布和组合。这个阶段是最难的，因为设计芯片虽然本质上是个搭积木的游戏，但是需要在很小的空间里搭上亿个半导体元件mos管，每个管子的尺寸达到了纳米量级，这难度就相当大了。顶级芯片厂商英特尔在美国的研发部门有上万人，其中有八千多个博士，可见，研发难度之大。目前，我们中国的芯片产业研发还集中在中低端领域，高端芯片研发基本为零。

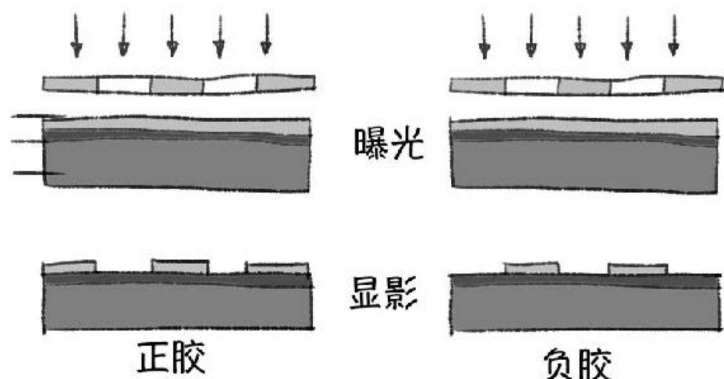
第二个阶段是制作。制作需要使用“光刻”的方法。因为尺寸太小，使用普通机械方法无法加工，必须使用紫外线照射方法进行加工，因此就称为光刻。目前世界上只有少数几个厂商能够制作光刻机，价格非常

昂贵，一台光刻机就要数亿美元，而且，由于货源稀少，即使是这个价格，依然供不应求。制作阶段有中国企业参与，例如台湾的“联发科”“台积电”。

具体来说，首先要将沙子融化还原。沙子就是二氧化硅，我们从中提取出单质硅，并且切成硅片。然后在硅片上涂上一层光刻胶。

下一步就是利用设计好的模板，使用紫外线照射涂有光刻胶的硅片。相当于把电路图画在硅片上。

照射好之后，再用化学溶液清洗。由于光刻胶的性质，被光照射的部分和没有被光照射的部分，其一会被洗掉。




之后，将杂质在硅片表面进行沉积。一部分没有光刻胶覆盖的部分就有了杂质，而被光刻胶保护的部分就没有杂质。随后洗去光刻胶，就形成了我们需要的一部分结构。

由于现代的芯片结构非常复杂，所以需要多次进行涂胶沉积操作，这样才能形成多层结构。

第三个阶段是封装与测试。芯片制作好了之后，需要切割、连接外部电路，同时对芯片的能力进行测试。这是芯片制作的收尾阶段，中国企业在这方面参与较多。

虽说芯片的原理大家都清楚，但是无论是设计，还是制作，都需要巨额的资金、人才和时间的投入。因为芯片实在太小了，空气中有些灰尘，工厂中有些振动，都会造成芯片无法使用。有人问，为什么我们能

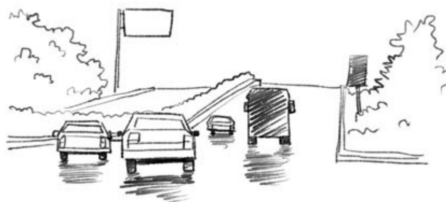
搞出原子弹、氢弹，能把太空飞船送上天，却造不出小小的芯片呢？实际上，芯片的确是人类智慧最顶尖的成果，难度就是比航天飞机还要大。

罗马不是一天建成的，英特尔等芯片厂商的技术优势也不是一天积累的。在芯片产业上，中国还有很长的路要走。

# 第三章　身边的科学

S—C—I—E—N—C—E

S



# 3.1　天空为什么是蓝色的？

——瑞利散射、太阳光的组成

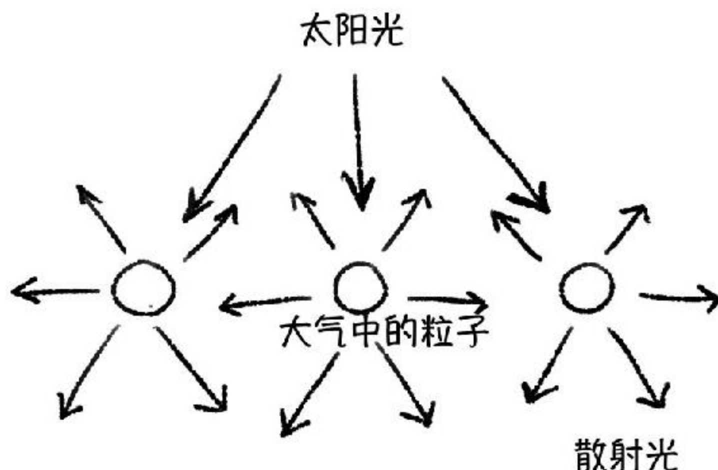
我们小的时候也许都问过家长，天空为什么是蓝色的？这个问题恐怕大多数家长都回答不出来。当我们长大了，变成家长了，孩子们又要问我们这个问题，这仿佛成了一个千古难题。在湖南卫视《天天向上》栏目里，汪涵是这样回答的：

天空是蓝色的，因为大海是蓝色的。大海是蓝色的，因为海里面有鱼，鱼会吐泡泡，blue、blue、blue，就把海染蓝了。

现在好了，看完这篇文章，你终于可以回答这个问题了。

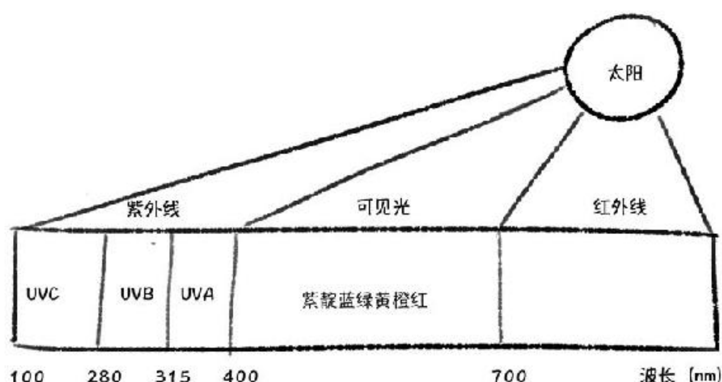
## 一、瑞利散射

我们知道，空气是由78%的氮气、21%的氧气和1%的其他气体构成的。空气中还存在一些杂质，例如固体颗粒、小水滴、小冰晶等。当光线射入大气层时，由于气体分子并不均匀以及杂质的影响，光线就会发生散射。所谓散射，就是将原本平行射入的光变为杂乱的方向。

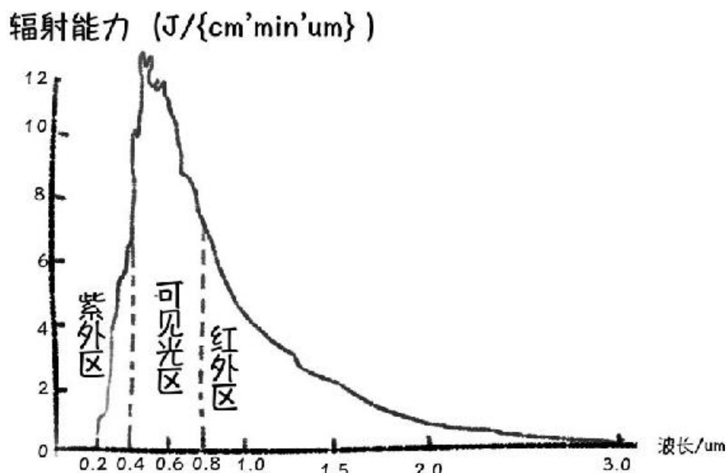


这样一来，原本向某个方向照射的太阳光，就会把整个天空照亮。如果没有大气的散射作用，就算在白天，天空中大部分还是黑的，就好像月亮表面上那样。

太阳光并不是单一波长的光，而是许多单色光合在一起构成的复色光。波长在400～760nm之间的光是可以被人眼看到的，称为可见光。太阳光的可见光部分可以分解为七种不同颜色，按波长从长到短依次是红、橙、黄、绿、蓝、靛、紫。除此之外，波长大于红光的不可见光称为红外线。波长比紫光短的不可见光称为紫外线。




太阳光的各种成分能量不是相同的。主要的能量都集中在可见光部分。这是很好理解的，人眼在进化过程中为了适应太阳，就把眼睛进化为对太阳光中能量最强的部分最敏感了。

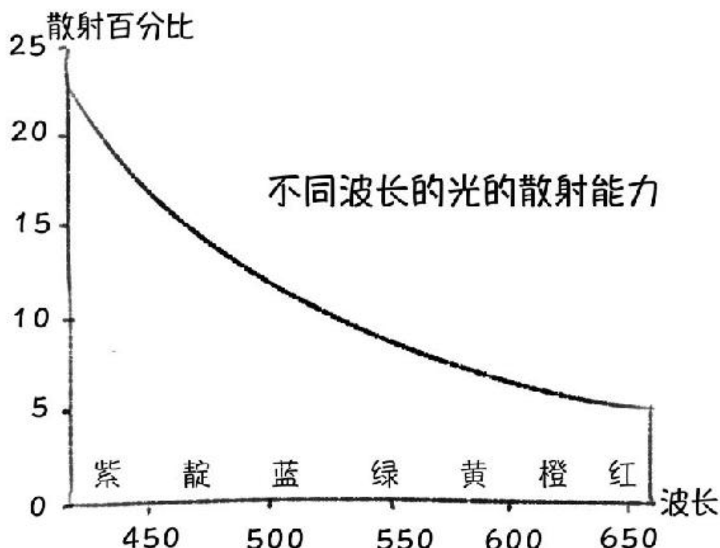



阳光在不同波长的辐射强度

不同波长的光散射能量也是不同的。卡文迪许实验室第二位主任，英国物理学家约翰·斯特拉特（第三代瑞利男爵）研究了光的散射问题，并得出了著名的瑞利散射公式。

瑞利指出：光被远远小于光波长的微小颗粒散射，散射光的强度与波长的四次方成反比。也就是说，波长越长，散射能力越差；波长越短，散射能力越强。

在可见光中，紫光散射能力最强，而红光的散射能力最弱。




同时，人的眼睛中有三种可以感觉颜色的视锥细胞，分别可以感受红色、绿色和蓝色，每种细胞对不同波长的光感受能力不同。相比来讲，人眼对黄绿色光最为敏感，这也恰好是太阳光中能量最集中的部分。

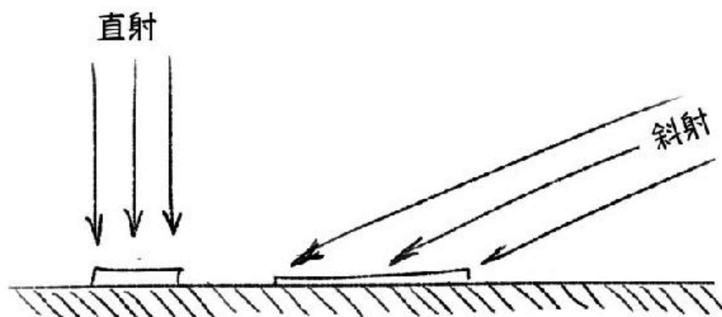
现在我们知道了：太阳光的能量分布在可见光最多，短波长散射能力最强，人眼对黄绿色光最敏感，三个因素叠加在一起，最终，人眼感受到天空的颜色就是蓝色了。不仅仅是天空，广阔的透明体，如水、玻璃等，也会呈蓝色。




## 二、朝阳和夕阳为什么是红色的？

利用这个原理，我们还可以解释许多问题。例如，早晨和傍晚的太阳光是偏红色的。

这是因为在中午的时候，太阳光垂直射向地面，所需要穿透的大气层比较薄。大部分的阳光都能到达地面，看起来太阳就是白色的。

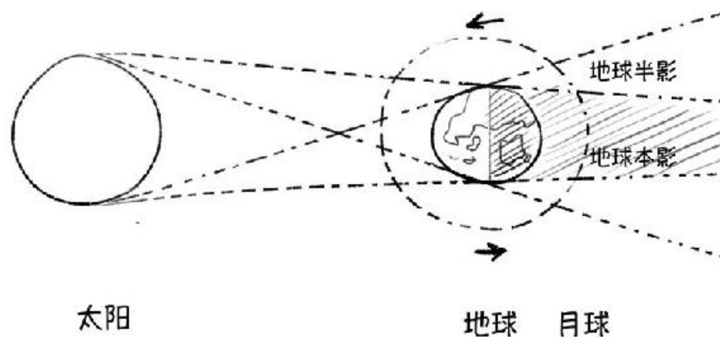



阳光的直射和斜射

但是早上和晚上，太阳光斜射向地面，在大气中经过的路程比较长。在这个过程中，短波长的光由于散射能力强，在经过地球大气的时候，大部分都散射掉了，余下的光就是散射能力最弱的长波长的光，于是就呈现出红色。

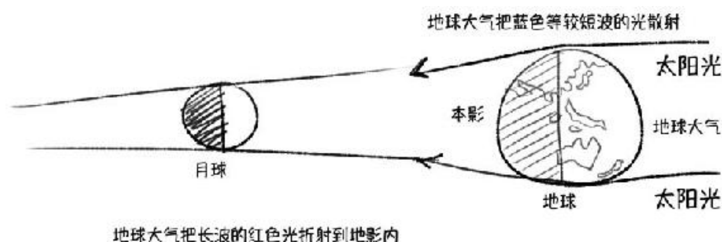
# 三、红月亮是如何形成的？

还有另一个很神奇的现象也可以用这个理论解释：红月亮。由于地球对太阳光的遮挡，在地球背面形成了本影区和半影区。在本影区，太阳光完全被遮挡，在半影区，一部分太阳光被地球遮挡。




如果月球处于地球的本影区里，就会出现月全食。如果月球一部分在本影区里，一部分在半影区里，就会出现月偏食。当月球扫过地球的本影区时，就会出现月偏食一月全食一月偏食的变化。

按理说，发生月全食时，月球应该消失不见，但事实上，月球会变成暗红色。这个原因是因为太阳光经过地球表面时，大气会把阳光进行折射，类似于一个凸透镜的作用，使原本接近平行的太阳光向内偏折，这样就能够照亮本影区中的月球。




但是，由于此时太阳光需要在大气中通过很长的路程，散射能力较强的短波长光都被散射掉了，能够通过的阳光主要是波长较长、散射能力较差的红光。所以，红光照射到月球上，就把月球照成了血月。红光散射能力差，穿透能力最强，所以人们的交通信号灯也采用红灯表示禁止通行。

看似简单的问题——天空为什么是蓝色的，其实并不简单，它蕴含了深刻的物理道理，这个问题直到100年前才真正被人们解释清楚。

# 3.2　星星为什么是黑白的？

——眼睛和视觉

我们夜晚遥望天空，会看到许多许多闪亮的星星。除了在地球附近的几颗行星之外，能够被我们看到的星星都是遥远的恒星。恒星就是像太阳一样的天体，它可以发光发热。

恒星距离我们非常遥远。比如，离我们最近的比邻星，距离我们有4.2光年。也就是说，它发出的光需要经过4.2年才能到达地球，我们此刻看到的比邻星，其实是它4.2年之前的样子。一些恒星距离地球有几十亿光年之远，我们看到的甚至可能是它们刚刚形成时的样子。

大家有没有发现，这漫天的星斗，肉眼看去基本都是黑白的，只有亮暗之分，却很难区分它们的颜色。那么恒星都是白色的吗？

## 一、恒星的颜色

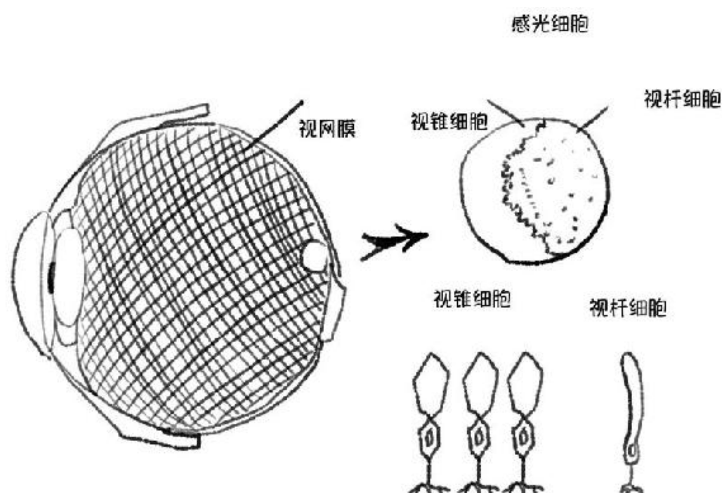
实际上，恒星的颜色大概可以分为红色、橙色、黄色、白色和蓝色，这是由于恒星表面的温度不同造成的，温度很低的恒星呈红色，而温度很高的恒星呈白色。

温度不同时，恒星的颜色为什么不同呢？这就回到了我们之前讨论过的一个黑体辐射问题：黑体会辐射各种波长的电磁波，如果温度越高，峰值（能量最高的电磁波）波长越短。

在可见光中，红光波长最长，紫光波长最短。当温度比较低时，恒星辐射的能量集中在长红光一侧，所以恒星表现为红色。当温度很高时，恒星辐射能量集中在紫光一侧，但是人的眼睛对紫光不敏感，而对邻近的蓝色光敏感，所以就表现为蓝色恒星。

# 二、视锥细胞和视杆细胞

既然如此，那么我们为什么看到的星空是黑白的呢？这又是人的眼睛结构所决定的。




人的眼睛相当于一个高级的照相机。光线经过角膜和晶状体，就好像通过照相机的镜头一样，光线会汇聚到后面的视网膜上。视网膜上有感光细胞，感光细胞感受到光照后，会把信息传给大脑。

感光细胞有两种：视锥细胞和视杆细胞。

每只眼睛大约有700万个视锥细胞，它们又可以分为三种，分别可以感受红色、绿色和蓝色。三种视锥细胞配合起来，就能感受到五颜六色的世界。不过，视锥细胞只能在强光下工作。

视杆细胞的作用刚好与视锥细胞互补。每只眼睛大约有1.2亿个视杆细胞，对光子的感受能力比视锥细胞强一百多倍，一个光子就可以激发视杆细胞的活动，所以在强光和弱光下都可以工作。但是，视杆细胞不能辨别颜色。

也就是说，只有在强光下，视锥细胞才能发挥作用，人们才能分辨颜色。在弱光下，只有视杆细胞才能发挥作用，因此人们是不能辨别颜色的。

许多动物的眼睛里都没有视锥细胞，便不能辨色，如猪、狗、牛等，其实都不能辨色，在它们的眼中，世界都是黑白的。西班牙斗牛士经常用一个红布斗牛，但实际上，牛对红色并没有什么特殊感觉，它只是讨厌一个人用抖动的布挑逗它而已，而那块布之所以是红色的，主要是为了让观众看得更清楚。



这种现象也是进化的结果，因为许多动物为了保护自己，夜间也需要清晰的视觉，所以对视杆细胞的需求要大于对视锥细胞的需求。相反，许多鸟类都有比人类更强的辨色能力，这同样是进化的结果，因为鸟类需要通过颜色辨别地面上的昆虫，从而进行捕食。

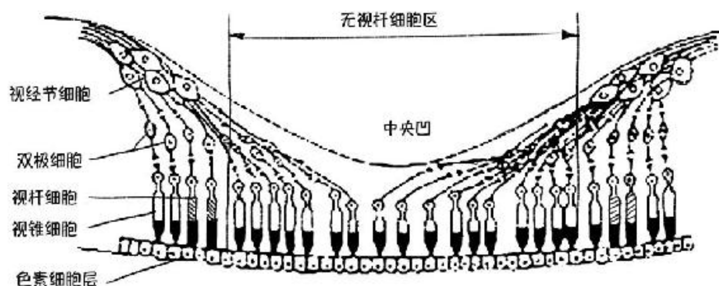
现在，我们就能解释肉眼看星星为什么只有黑白的了。因为遥远的恒星发射的光到达地球时，光线已经很弱了。所以，只有视杆细胞才能感受到这些星光，视锥细胞是不能工作的。而视杆细胞又无法分辨颜色，所以看到的星星就只是黑白的了。




# 三、一个有趣的光学实验：盲点

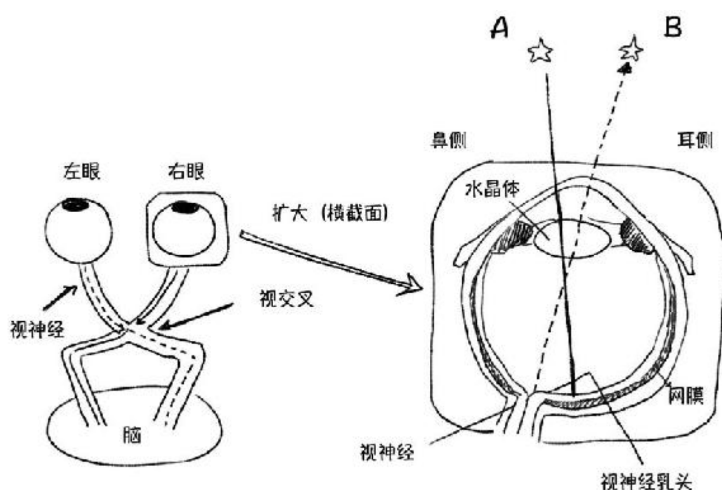
但你有没有发现：如果我们盯着一颗很暗的星星看，它就会消失不见。如果我们把目光移向别处，它却又冒出来了，这是为什么呢？

这是因为视锥细胞和视杆细胞并不是均匀分布在视网膜上的。在中央凹附近，视锥细胞是最密集的，所以可以最清楚地分辨物体的颜色。当我们注视一个点，角膜和晶状体就会将这个物体的像呈在这个位置。但是，这个位置是没有视杆细胞的。所以，如果一个物体发射的光太暗，只有视杆细胞能看到它，我们盯着它看，将它成像在没有视杆细胞的中央凹处，于是它就消失不见了。




关于眼睛，还有一个很神奇的现象：盲点。

由于感光细胞必须经过视神经连接到大脑之中，视神经在视网膜上有一个节点，称为视神经乳头，这个地方是没有感光细胞的。如果物体的像呈在这个部位，就没有办法被大脑感知到。




找盲点的方法也很简单：在一张纸上相距10cm左右画两个小五角星$A$ 和$B$ ，闭上左眼，用右眼观察左边的五角星$A$ ，此时$A$ 会成像在中央凹处。前后移动头，在某个合适的位置，$B$ 会成像在神经节处，此时余光就无法看到$B$ 了。

眼睛是一个非常神奇的器官，物体发射的光线进入了我们的眼睛，才让我们感受到这个丰富多彩的世界。



# 3.3　正常夫妻为啥会生出色盲孩子？

——遗传：显性基因和隐性基因

我们之前讲过，眼睛中有两种感光细胞：视锥细胞能分辨颜色，视杆细胞不能分辨颜色，但是感光能力强。鱼类、鸟类和爬行动物的视网膜上都有大量的视锥细胞，因此可以辨别颜色，哺乳动物中有辨色能力的动物并不多，但是夜视能力却比较好。灵长类动物能够辨色，当然，也包含了人类。

但是，并不是所有人的辨色能力都是相同的。英国化学家、近代原子论的创立者道尔顿是色盲的发现者，所以色盲症也被称为“道尔顿症”。

有一年的圣诞节，道尔顿为母亲买了一双深蓝色的袜子以表示自己对老人的孝敬。当他送给母亲时，母亲却厉声责问他，为什么买一双红色袜子。依照当地宗教习俗，妇女禁忌红色。由此道尔顿才发现自己的辨色能力与众不同。他经过认真的调查，发现他哥哥也和他一样，具有不正常的辨色能力，另有一些人也具有这一病症。1798年，他出版了第一部论述此问题的科学专著《关于色彩视觉的离奇事实》。



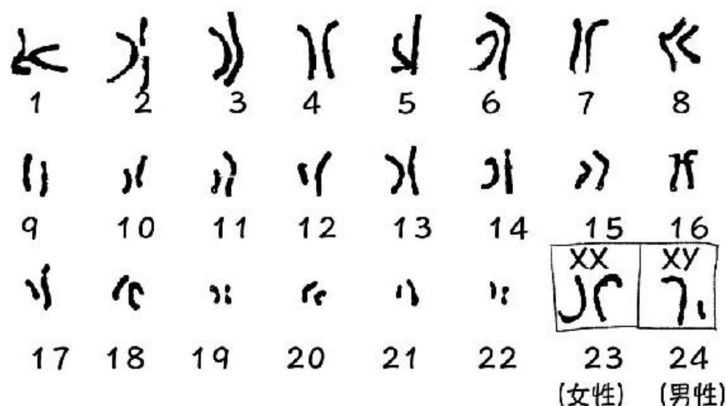

现在我们认识到，色盲是一种很常见的遗传问题，并不能称之为疾病。在某些特定的情况下，色盲人群会比普通人群更加适应环境。红绿

色盲者非不能分辨任何颜色，只是对于某些颜色的感受与常人不同。如果一个人不能分辨任何颜色，只能看到黑白的世界，就称为全色盲，全色盲在世界上比较少见。

色盲到底是如何形成的呢？

# 一、孩子的性别是如何决定的？

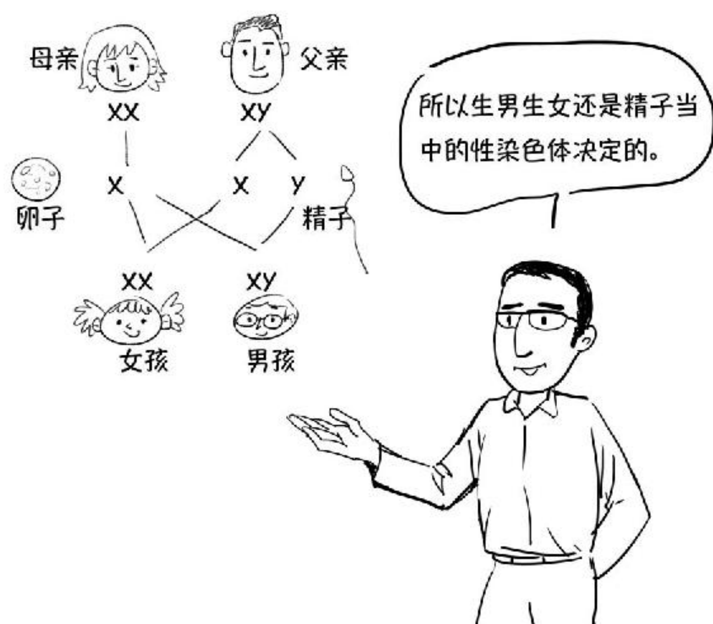
为了解释这个问题，我们首先要解释一下基本的遗传学原理。我们知道，人的某种特性都是由基因所决定的，而基因的载体是染色体。人体有23对46条染色体。其中22对染色体为常染色体，另一对染色体称为性染色体。性染色体决定了人的性别。




女性的两条性染色体大小形态都相同，称为X染色体。男性的性染色体中，一条是与女性相同的X染色体，另一条则小得多，称为Y染色体。于是常用“XX”表示女性，用“XY”表示男性。

性成熟时，男性会产生精子，精子中只有23条染色体，是正常体细胞染色体数目的一半，它实际上是在人体的23对染色体中每对随机抽取一条形成的，那么第23条染色体既可能是X染色体，也可能是Y染色体。同样，女性产生卵子，也含有23条染色体，其中，第23条染色体一定是X染色体。当精子与卵子结合形成受精卵时，对应的染色体结合，也就是说，孩子的染色体一半来源于父亲，一半来源于母亲。尤其是第23对

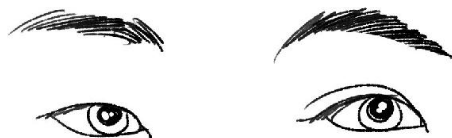
染色体，如果是两条X染色体结合，就会生出女孩，如果是一条X染色体和一条Y染色体结合，就会生出男孩。




# 二、显性和隐性基因

染色体上的各种基因决定了生物体的特征，基因在每对染色体上对应的位置成对存在，其中一个影响力大，能够决定后代的性状表现，称为显性基因；另一个影响力小，不能单独决定后代的性状表现，称为隐性基因。

我们常用A表示显性基因，a表示隐性基因。一对基因就有AA、Aa、aA和aa四种可能。只要含有显性基因，生物体就表现出显性特征，比如AA、Aa和aA。只有一对基因都是隐性基因aa时，生物体才会表现出隐性特征。比如双眼皮就是显性特征，而单眼皮就是隐性特征。



如果某些遗传特征的基因是位于性染色体上，就称为伴性遗传。色

盲是典型的伴性遗传特征。控制色觉的基因位于性染色体中的X染色体上，色觉正常是显性特征，色觉异常是隐性特征。

我们用A表示正常色觉基因，用a表示色盲基因，男性只有一条X染色体，所以色觉基因携带情况只有$X^A$ Y和$X^a$ Y两种，$X^A$ Y是显性特征：色觉正常，$X^a$ Y表现为隐性特征——色盲。

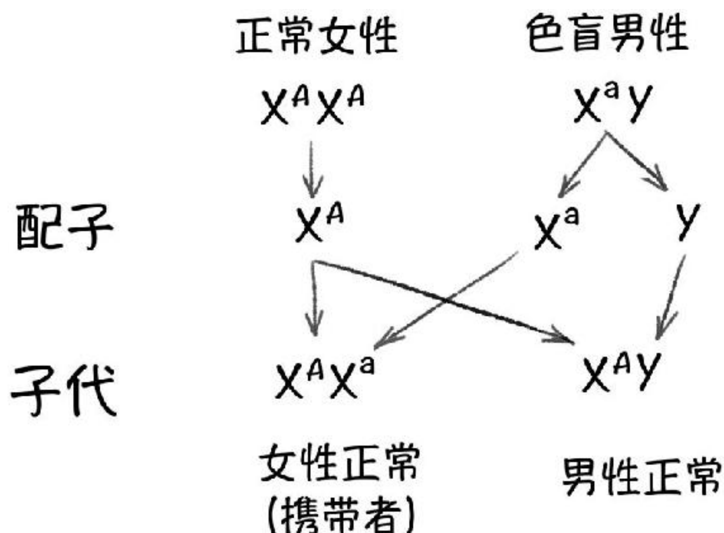
女性有两条X染色体，所以色觉基因携带情况有四种：$X^A$ $X^A$ 、$X^A$ $X^a$ 、$X^a$ $X^A$ 、$X^a$ $X^a$ ，其中，前三种由于有显性基因A，表现为色觉正常；第四种两个色觉基因都是隐性，表现为色盲。

由于女性有两条X染色体，只要有一条含有显性色觉基因A，就不是色盲，而男性只有一条X染色体，如果携带色盲基因a，就是色盲，所以男性色盲率远高于女性。全世界范围的男性中，红绿色盲约占8%，而女性中红绿色盲约占0.5%。色盲概率如此之高，与左利手的概率相差无几。如果在一个班级中有50名同学，其中，有色盲的概率大约是90%，所以色盲是一种非常常见的遗传特征。

# 三、伴性遗传规律

那么，父母如果是色盲，孩子就一定是色盲吗？我们不妨来分析两种典型情况：

比如，如果一个完全正常的女性（$X^A$ $X^A$ ）和一个色盲男性（$X^a$ Y）结合，孩子的性染色体中的一条来源于母亲，另一条来源于父亲。母亲一定给出$X^A$ 染色体，而父亲可能给出$X^a$ 染色体，也可能给出Y染色体。

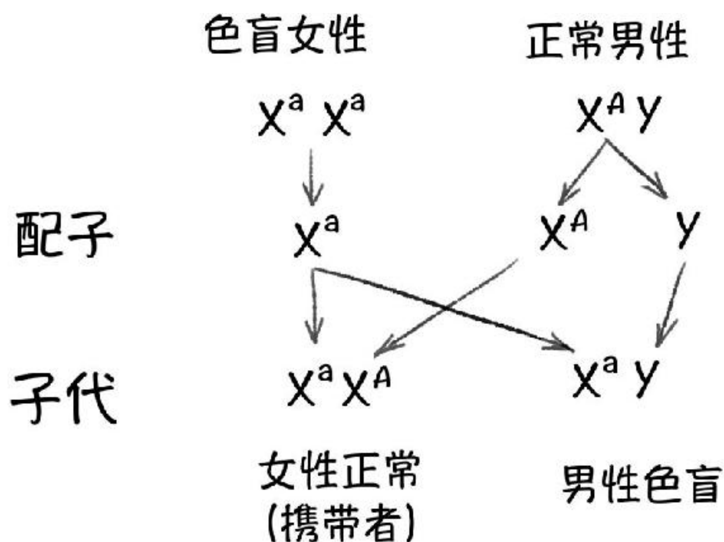



如果生的孩子是女孩，那么就是$X^A$ $X^a$ ，这种情况下，孩子的表现是正常的，但是是色盲基因携带者，她可能会把这条色盲基因传给自己的下一代。

如果生的孩子是男孩，那么就是$X^A$ Y，这种情况下，孩子不光表现是正常的，而且父亲的色盲基因被抛弃了，孩子的色觉基因也是完全正常的。

也就是说，如果一个完全正常的女性和色盲男性结合，生的孩子一定是表现正常的。但是如果是女孩，将会是一个色盲基因携带者。

那么，如果一个色盲女性（$X^a$ $X^a$ ）和正常男性（$X^A$ Y）结合，情况又是如何呢？母亲给出一条性染色体，必然是$X^a$ ，男性给出的性染色体可能是$X^A$ ，也可能是Y。

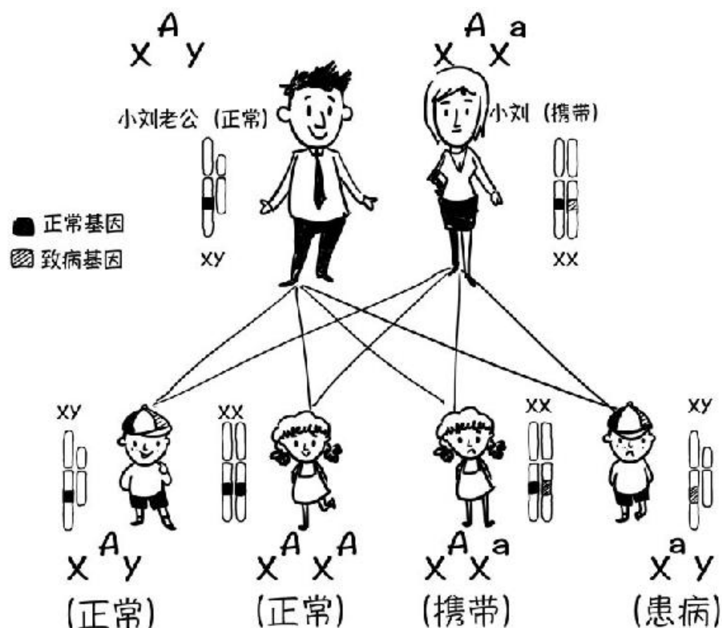



如果生的孩子是女孩，那么性染色体是$X^a$ $X^A$ 此时女孩表现是正常的，但是是色盲基因携带者。

如果生的孩子是男孩，那么性染色体就是$X^a$ Y，此时男孩只有一个a的色盲隐性基因，因此表现为色盲。

也就是说，如果妈妈是色盲，则儿子一定是色盲。

还有同学问：为什么我的父母都是正常的，而我是色盲呢？这是因为你的母亲一定是色盲基因携带者$X^A$ $X^a$ ，而父亲是正常的$X^A$ Y。此时如果生出女儿，性染色体情况为$X^A$ $X^A$ 或$X^a$ $X^A$ ，都表现为正常，生出儿子是$X^A$ Y或者$X^a$ Y，其中，$X^A$ Y是正常，而$X^a$ Y表现为色盲。也就是说，妈妈是色盲基因携带者，孩子有1/4的概率是色盲。如果是儿子，则有1/2的概率是色盲。

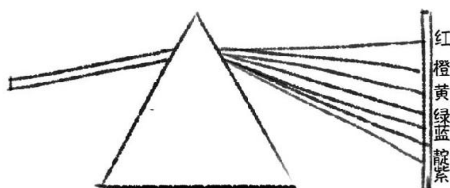



绝大多数的遗传缺陷在生物进化的过程中都被淘汰了，但是色盲能够长期保持下来，就是因为它是最不影响生活的遗传缺陷。而且，在某些特定的情况下，比如暗光等情况下，色盲者可能会比常人具有更好的视力。只是，由于有些专业需要辨色能力，所以色盲的同学在高考报考时不能报考化学、医学、生物、交通等相关专业，不能辨别红绿灯的色盲人群也无法取得驾驶执照，给生活带来了一定的不便。

# 3.4　双层彩虹是怎么回事？

——色散的原理

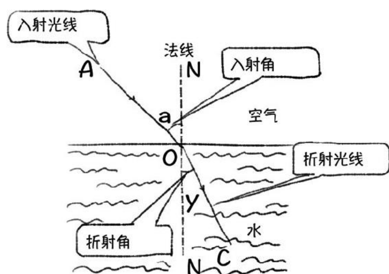
雨过天晴，我们会看到美丽的彩虹。有时候，彩虹是两条一起出现的。如果大家仔细观察，就会发现较低的彩虹外侧是红色，内层是紫色，称为“主虹”或者“虹”，而较高的彩虹外侧是紫色，内侧是红色，称为“副虹”或“霓”，所以“霓虹”才是彩虹的全称。而这种现象是如何形成的呢？




## 一、光的折射

为了解释彩虹的形成，我们首先要解释光的折射。

当光从一种介质斜射入另一种介质时，传播方向一般会发生变化。例如，光线从空气斜射入水中时，折射光线会更偏向垂直水面的轴线，这个轴线称为“法线”。



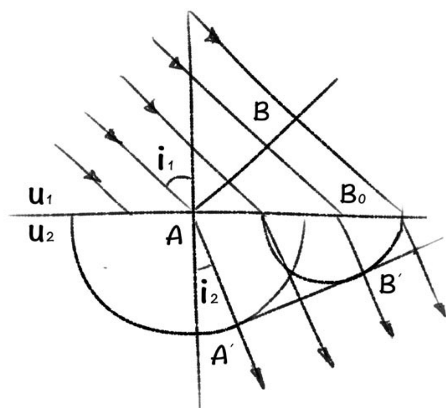




最早对折射现象进行正确解释的人是荷兰物理学家惠更斯。



他指出：波发生折射的原因是波在不同介质中的传播速度不同。并且，根据光的折射规律，光在空气中传播速度快，在水中传播速度慢。



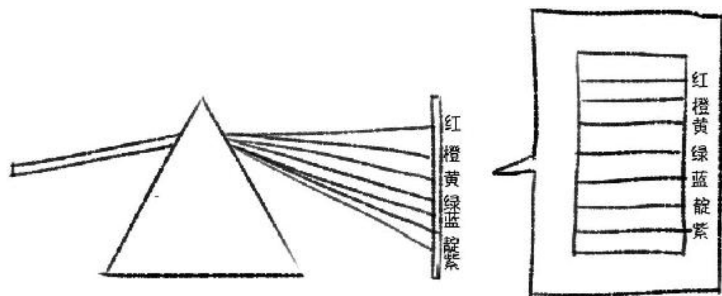

如上图所示，一束平行光照射到两种介质的界面时，各条光线并不是同时到达界面的。下方的光线先到界面$A$ 点，上方光线此时只到达$B$。随后，$A$ 就会在第二种介质中传播，而光线$B$ 继续在第一种介质中传播。由于在第二种介质中光的传播速度比较慢，当$B$ 光线从$B$ 传播到界面上的$B_0$ 点时，光线$A$ 会从$A$ 传播到$A'$点，并且$AA' < BB_0$，这就造成了光线的偏折。宛如一辆汽车本来在柏油马路上走，右前轮突然进入了沙土地面，此时汽车的运动方向就会发生偏折。

# 二、光的色散

光在介质中的速度主要由介质所决定，例如水中的光速大约是真空

中的3/4，玻璃中的光速大约是真空中的2/3。但是，即便是同一种介质，不同颜色的光传播速度也有微小的差别，这就造成了不同颜色的光在折射时的偏折程度不同。

我们知道，白色光是由红、橙、黄、绿、蓝、靛、紫七种颜色的单色光构成的。如果一束白光进入介质，由于每种色光偏折程度不同，出射时就不再是白色，而是分成了七种颜色。这个现象就称为"色散"，最早是由牛顿发现的。



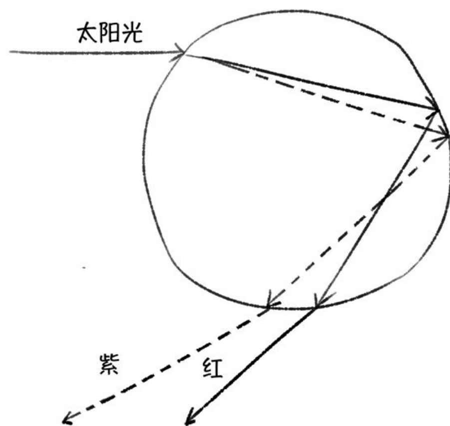

如上图所示，牛顿让一束白光通过三棱镜，他发现：射出光分为七种颜色。相比于原来的入射方向，红色光偏折最小，紫光偏折最大。这个规律在任何介质折射时都是适用的。

当空气中有小水滴时，光线进入小水滴会发生折射和反射。由于刚才所说的原因，不同颜色的光在发生折射时偏折程度不同，不同颜色的光彼此分开，人们就观察到了五颜六色的彩虹。

# 三、虹的形成

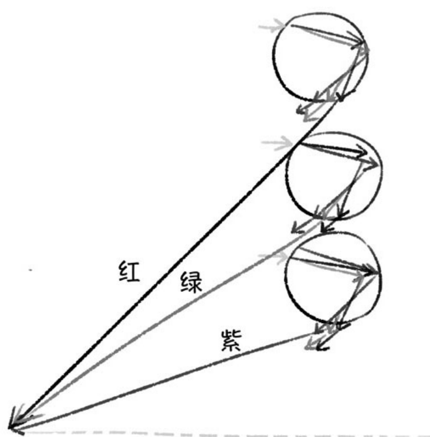
我们首先来解释一下主虹（虹）的形成。

光线进入水滴时，可能会发生两次折射和一次反射，然后射出水滴，如下图所示。




光线首先通过一次折射进入水滴，再通过一次反射和一次折射射出水滴。由于色散的原因，对于同一个水滴的出射光，紫色光偏折程度大，偏向水平；而红色光偏折程度小，偏向竖直。于是就形成了紫光在上、红光在下的情况。在红光到紫光之间，分布着橙、黄、绿、蓝、靛几种颜色。

当人们进行观察时，空气中分布的许多水滴都会形成类似的效果，如图所示。

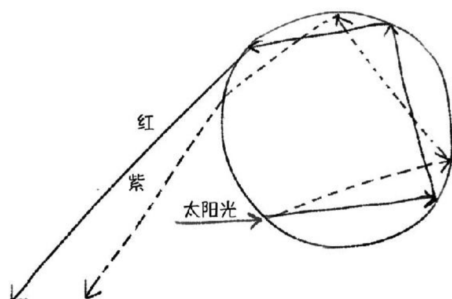



大家注意：同一个水滴发出的光不能全部进入我们的眼睛，因为一个水滴色散后的光线是向不同方向发射的。我们所接收的是不同水滴射出的光。这样一来，接收到的最上方的光更加偏向竖直，相比于原来的入射方向偏角最小，是红色；最下方的光线更加偏向水平，相比于原来光线的入射方向偏角最大，因此是紫色。于是我们就看到了外红内紫的

主虹。

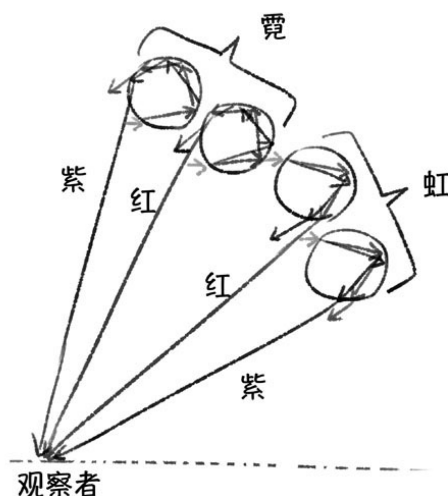
# 四、霓的形成

副虹（霓）的形成原因类似，只不过，它是光线在水滴中进行了两次折射和两次反射形成的。




如图所示，在这种情况下，由于红光偏折程度小，射出时更加偏向水平；紫光偏折程度大，射出时更加偏向竖直，刚好与主虹相反，于是就造成了外紫内红的效果。

一般而言，虹与霓会同时出现，而且霓更高。只是由于霓多了一次反射，比虹弱很多，所以有时候人们认为只有一条彩虹。




美丽的彩虹让人们产生了诸多幻想，古希腊人认为彩虹女神伊里丝

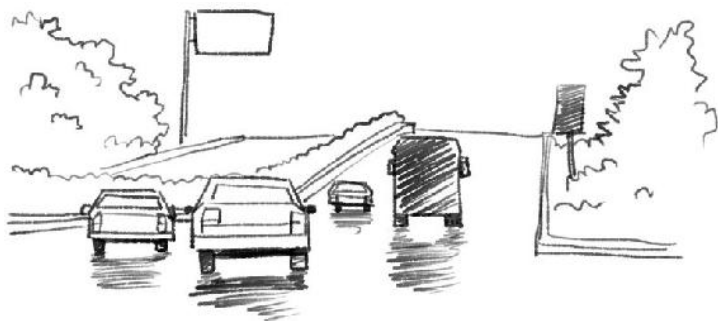
是宙斯的使者。但是只有科学才能让我们真正认识自然，了解自然。



# 3.5　炎热的夏天为啥总感觉马路上有水？

——海市蜃楼的原理

不知道大家在炎热的夏天，有没有看到过这样一个现象：在不远的前方，路面上好像有一滩水，还映出了前车的倒影。但我们走过去一看，其实路面上根本就没有水。



你并没有出现幻觉，这是一种光的折射和全反射现象，通常称作“海市蜃楼”。

## 一、斯涅尔定律

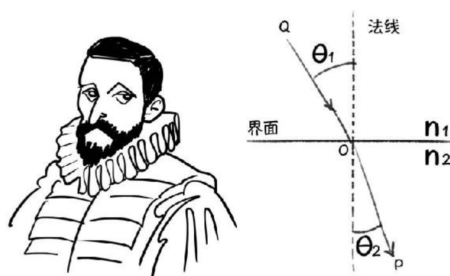
我们在上一节说过，光在不同介质间传播的时候，传播方向一般会发生变化。比如一束光从空气斜射入水中，会向下偏折。出现折射的原因是因为光在不同介质中的传播速度不同。人们用折射率$n$ 表示光在不同介质中传播速度的不同，折射率$n$ 等于真空中光速$c$ 与介质中光速$v$ 的比：$n=\frac{c}{v}$ ，也就是说，介质的折射率$n$ 越大，光在介质中传播的速度$v$ 越小。

显然，真空中光速$v=c$ ，折射率$n=1$。空气中光速接近于真空中光速，因此，其他介质中光速都小于真空中光速，折射率$n>1$。

| 介质 | 折射率 |
| --- | --- |
| 空气 | 1.0003 |
| 冰 | 1.309 |
| 水（20℃） | 1.333 |
| 普通酒精 | 1.360 |
| 面粉 | 1.434 |
| 玻璃 | 1.500 |
| 翡翠 | 1.570 |
| 红宝石 | 1.770 |
| 水晶 | 2.000 |
| 钻石 | 2.471 |

折射率主要由介质材料所决定，但是与光的波长也有一定关系，这就是上面说的色散现象。上表中的折射率指的是这种材料对光的平均折射率。

光线在折射率不同的两种材料中传播时，角度的关系满足斯涅尔定律，用荷兰物理学家威理博·斯涅尔的名字命名。



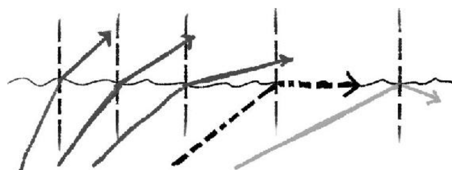

$$n_1 \sin\theta_1 = n_2 \sin\theta_2$$

其中，$n$ 是折射率，$\theta$ 是介质中光线与法线的夹角，$\sin\theta$ 是正弦函数，它在0～90°度的区间内是一个增函数，也就是说，角度$\theta$ 越大，$\sin\theta$ 也会越大。通过这个公式我们就会发现，如果材料的折射率$n$ 比较大，那么介质中的角度$\theta$ 就小，这种介质就称为“光密介质”；如果材料的折射率$n$ 比较小，那么介质中的角度$\theta$ 就大，这种介质就称为“光疏介质”。

比如，空气相对于水就是光疏介质，水相对于空气就是光密介质。光从空气射入水中，折射率变大，折射角小于入射角，因此光线向法线偏折。

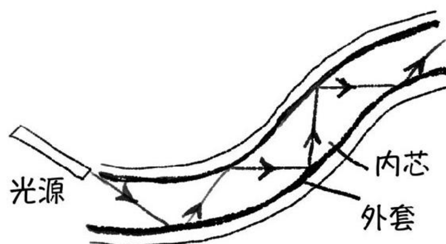
# 二、全反射现象

如果光线从水中射入空气中，情况又是如何呢？由于水是光密介质，角度较小，空气是光疏介质，角度较大，因此光线会向水面偏折。如果增大水中光线的入射角，那么空气中的折射角也会增大，在某个时刻，空气中的折射角达到90°，此时就称为“掠出射”。如果继续增大入射角，那么折射光线无论折射向哪里，都不合理了。此时会出现一种奇特的现象，折射光线消失，只剩下反射光线，这种现象就称为“全反射”。



光线以不同角度射出水面

全反射在生活中有很多应用，比如光导纤维就是利用光线在内芯和外套之间反复全反射传播信号的。




# 三、海市蜃楼

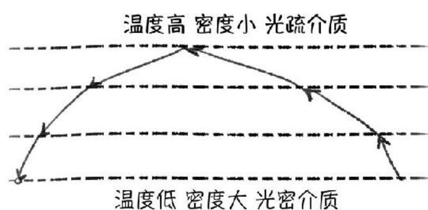



现在我们就可以解释“海市蜃楼”了。先说说什么叫“蜃”呢？

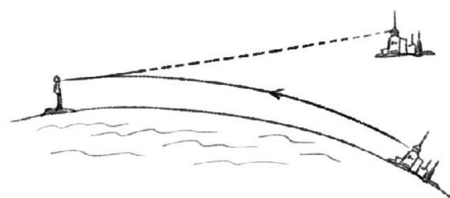
龙生九子，各不相同，其中有一个儿子就叫蜃，蜃喜欢吞云吐雾，它在海上把东西都吞到肚子里去，一会儿又吐出来，这时，人们就可以看到有些东西浮在海平面上。这只是一种神话传说，“海市蜃楼”实际上是一种光的折射现象，在海洋和沙漠中都可能出现。



在海洋上，海水比热容更大，也就是说，在接受太阳照射时，海水不容易升温。在强烈的太阳照射下，靠近海水的地方，空气温度比较低，密度较大，折射率较大，属于光密介质。高层空气温度较高，空气受热膨胀，密度较小，折射率较小，是光疏介质。光线从光密介质射向光疏介质，就像从水中射向空气一样，折射角变大，光线会趋于水平。假如海面上有一艘船，这艘船反射出的光线向上射，就会在各个不同的空气层之间发生折射，发生弯折。

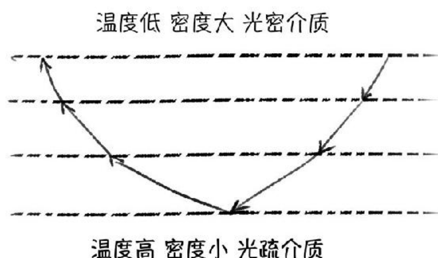



如果在还没有发生全反射时，光线就进入了人眼。那么人眼就会以为物体在远方的高处，形成正立的蜃景。如果光线在传播过程中发生了全反射，人们逆着光线看去，就会看到倒立的蜃景。



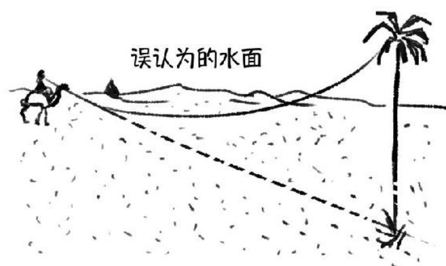
海洋上的海市蜃楼

沙漠中的海市蜃楼成因刚好与海洋上相反。沙子的比热容很小，所以靠近沙漠的地方，空气温度比较高。空气受热膨胀，密度较小，折射率较小，是光疏介质。上层的空气距离沙子较远，温度相对较低，空气密度相对较大，折射率较大，也就形成了光密介质。光线从上向下照射时，从光密介质进入光疏介质，相当于从水射向空气，折射角变大，光线偏向水平。当折射角增大到一定程度时，就会发生全反射而向上照射。




比如有一朵云彩飘在空中，它反射的光线经过折射和全反射被地面上的人观察到。大脑会认为光线依然是直线传播的，因此认为云彩在地下。云彩不会在地下，所以人们会认为地面上有一个可以反射光线的物质，那就是水，这就是沙漠中海市蜃楼的原理。由于沙漠中沙子的温度非常高，形成蜃景的光线在靠近沙子时一定会发生全反射，所以在沙漠

中的海市蜃楼都是倒像。




沙漠中的海市蜃楼

炎热的夏天，马路上的温度也非常高，形成的效果与沙漠相同。所以在靠近地面的位置，光线很容易发生全反射，映出车辆的倒影，所以就会让人们误以为地面上有水。下次再遇到这种现象，别再被你的眼睛欺骗了哦。

# 3.6　雨中走路淋雨多，还是跑步淋雨多？

——数学模型

如果有一天下雨了，你刚好没有带伞，还没有地方避雨，那么你会选择在雨中漫步，还是奔跑呢？

这是一个古老的问题，引起了国内外的数次讨论。中央电视台《加油向未来》节目还做了个实验，得出了跑步淋雨少的结论。国外的节目《流言终结者》也做了实验，而且做了两次，居然得到了相反的结果。

原因在于，这个问题在实际中影响因素非常多，例如雨量、风速、人的速度、人的表面积和形状，等等，都会影响实验结果。尤其是如果雨滴下落时不均匀，那么随机性就会变得更大。

## 一、物理模型

我在这里基于简单的物理模型做一个分析。所谓物理模型，就是从一个实际的复杂问题中抽象出最核心的内容，而忽略其他不重要的影响因素。例如我们研究地球围绕太阳运动，就把地球看作一个点，而不去管地面上的山川河流，这就是质点模型。作为一个模型，我们必须给出一些假设，虽然这些假设可能与实际情况并不完全相同。

假设1：雨是均匀的，雨滴无限小，而且各处密度均匀，单位体积内雨的质量为ρ 。

假设2：没有风，雨滴匀速下落，速度为$v$ 。

假设3：人匀速运动，运动的速度为$u$ 。

假设4：把人的身材看作一个长方体，人身体前方的面积为$S_1$ ，头

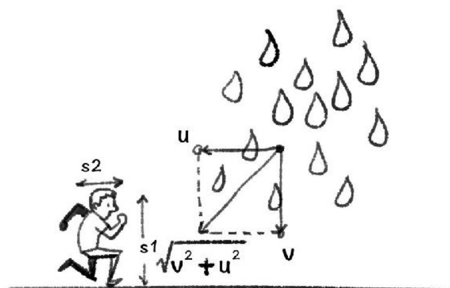
顶的面积为$S_2$。

假设5：人的目标是从A地到达相距$L$的B地。

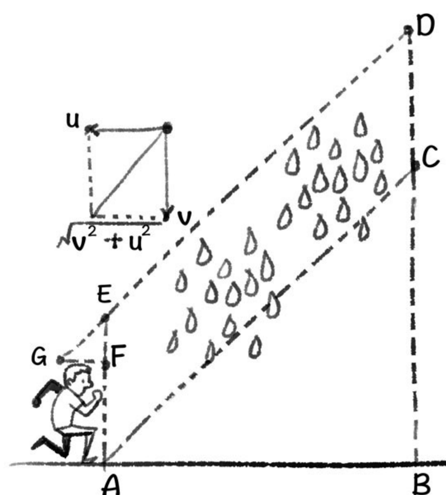
有了以上假设，我们就可以进行计算了。人在雨中向前运动，头顶会淋雨，前面也会淋到雨。我们要分别计算这两部分。

# 二、基本分析

我们首先要研究的是空间中哪些雨落在了人的身上。如果选取地面作为参考系的话，人在运动，雨也在运动，问题会比较复杂。我们可以换一个参考系——以人为参考。这样一来，人就可以看作静止不动的，而雨滴在竖直方向具有下落的速度$v$，在水平方向具有相对于人向后的水平速度$u$，雨滴相对于人就是斜向下运动的，如下图所示：




人在从A地到B地的过程中，雨滴相对于人向斜下方匀速直线运动，能够落到人身上的雨滴（忽略人头顶的一个小三角形）都在他斜前方一个柱体内，如下图中的ACDE部分。这些雨滴会朝着人奔跑，最终撞到人身上。

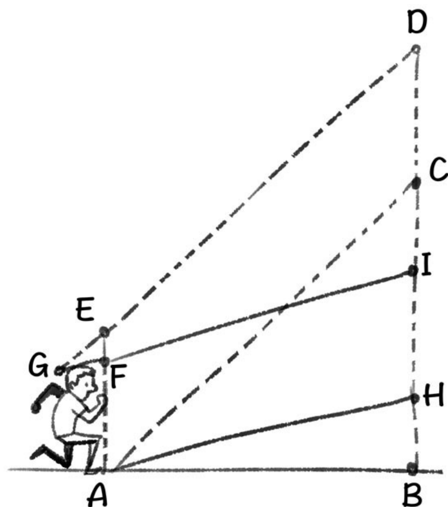



这是一个斜柱体，它的底面积是人迎接雨滴的截面积$S$，如图中AE部分所示。而柱体的高是AB之间的距离AB＝$L$。根据柱体体积公式得到雨滴体积$V=SL$，单位体积的雨滴质量为ρ，于是最终落到人身上的雨滴的总量为：$m=pSL$

# 三、如何淋雨少？

那么，如何才能淋雨更少呢？

显而易见，无论以多大速度奔跑，AB之间的距离$L$是一定的。当奔跑速度不同时，雨滴相对于人的速度不同，因而柱体的倾斜程度不同，截面积$S$也不同。



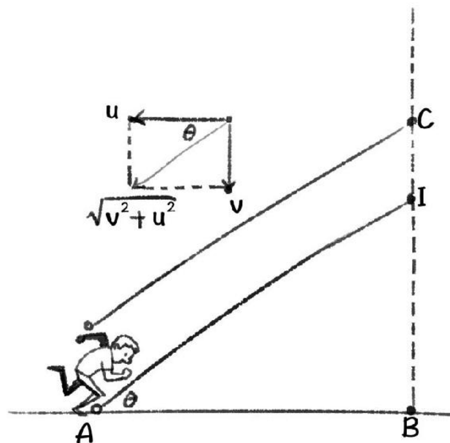

如上图所示，如果人的奔跑速度比较大，雨滴相对于人速度更接近水平，这样人迎接雨滴的截面积为AF部分；如果人的奔跑速度比较小，雨滴相对于人速度方向更加竖直，人迎接雨滴的面积是AE部分。

显然，柱体AFIH和柱体AEDC的高相同，但是AF部分面积更小，柱体体积更小，柱体中的雨质量更小，即人以更大速度奔跑时，淋雨少。如果人以无限大的速度奔跑，则雨滴一点也不落到头顶，而是全部落在人的身体前侧面。

# 四、还能再给力一点吗？

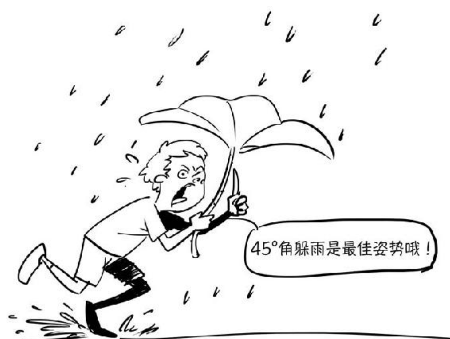
如果人已经达到最大奔跑速度了，那还有没有可能继续减少淋雨呢？

其实我们还有办法。因为人的头顶面积小于身体前面的面积，我们可以让身体倾斜过来迎接雨滴，这样就可以使得人迎接雨滴的面积进一步减小，雨柱体变得更细。




如果要获得最小的迎接雨的底面积，就应该完全用头顶面积迎接雨，此时人的倾斜程度应该与雨相对于人的速度方向平行。如上图所示，这个夹角用三角函数表示就是：$\tan\theta=\frac{v}{u}$。

比如人的奔跑速度和雨滴下落的速度相同时，人向前倾斜45°角是最好的。




综上所述，在一定的模型条件下，人以尽量大的速度奔跑，并且使身体向前倾斜，可以使落到身上的雨滴减少。如果我们可以精巧地调整身体的角度，使得总是只有头顶迎接雨滴，那么我们只需要一小块挡住脑袋的荷叶，就可以保证身上一点水都没有。

# 3.7 电磁炉是怎么加热食物的？

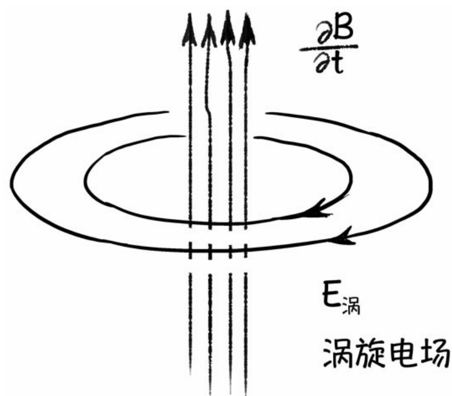
——涡流的产生与应用

电磁炉是一种常见炊具，它是使用电磁感应所产生的涡流对食物进行加热的。相比于传统的电炉和燃气灶，电磁炉具有很多优势，如热效率更高、炉面清洁、不产生有毒物质一氧化碳、火力稳定且易于控制等。

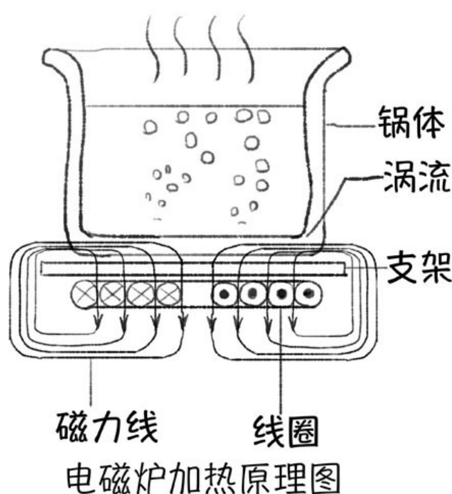


## 一、涡流

电磁炉具体是如何工作的呢？之前我们说过，英国物理学家法拉第最早发现了电磁感应现象，即当磁场变化时，导体中有感应电流产生。随后，科学家麦克斯韦推断，电磁感应产生的原因是变化的磁场会在周围空间产生电场，这种电场与磁场垂直，并且首尾相接，称为“涡旋电场”。如果在涡旋电场处存在导体，电场就会推动电荷运动，产生涡旋电流。




电磁炉就是利用这种原理制作的。在电磁炉内部，通过一定的方法，将50Hz的常用交流电变为直流，然后再变为20kHz左右的高频电流。高频电流通入电磁炉面板中的线圈里，就会产生高频磁场。



电磁炉加热原理图

线圈产生的高频磁场会在周围产生高频电场，这个电场遇到导体——锅底，就会产生涡旋电流。由于铁锅存在电阻，因此这个涡流就会产生热量。也就是说，电磁炉的加热原理是直接在锅底产生电流加热食物，而电磁炉面板发热量很少。锅底的热量多数被食物所吸收，因此热效率很高。据统计，电磁炉的能量转换效率达90%，传统电热炉为70%，而燃气灶因为加热了周围空气，热效率只有30%。如果电磁炉打开，但是不放铁锅时，涡旋电场附近没有导体，因此就不会产生电流，也不生热。此时电磁炉会报警，同时断电，相比于传统炉具更加安全。




电磁炉一般要使用铁锅或者不锈钢锅，这是因为铁锅有两个好处：

第一，铁锅是导电的。如果用陶瓷锅，其不导电，就不能产生涡流。

第二，铁锅是铁磁性的。所谓铁磁性，就是在外加磁场的情况下，物质会被磁化，形成与外界磁场相同方向的磁场，从而可以加强外界磁场。高频电流产生的磁场被加强之后，可以产生更强的涡旋电场，这样，电流才足够大。如果使用铜锅或者铝锅，因为这两种金属不是铁磁性的，不能加强磁场，因此涡旋电场不够大。

## 二、电磁炉对人体有危害吗？

由于电磁炉产生高频振荡磁场，会有电磁波产生。但是，电磁炉的频率只有2万Hz左右，相比于微波炉20亿Hz，频率低了十万倍。而且，电磁炉辐射电磁波的范围很小，只局限在电磁炉附近一个很小的范围内，强度也不足以对普通人产生危害。




不过，这种电磁波依然可能对某些电子设备产生影响，例如装有心脏起搏器的人，最好避免距离电磁炉太近。

# 三、涡流的应用和危害

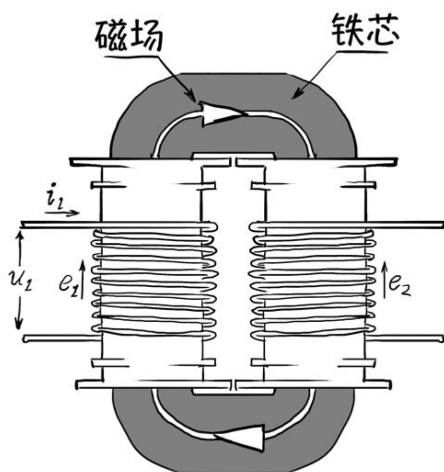
除了微波炉，涡流在生活中还有许多应用。例如，机场安检时使用的手持式安检仪（金属探测器）的原理与电磁炉相同。



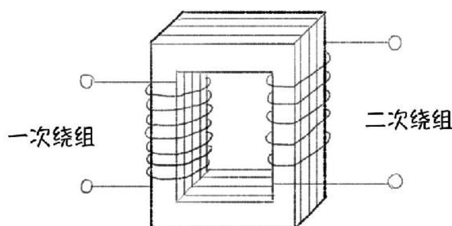

这种安检仪内部有个线圈，通过交流电时会产生变化磁场，变化磁场又产生涡旋电场。如果探测器附近有金属导体，那么导体中就会产生涡流。这个涡流也会产生感应磁场，而探测器中有元件可以感受到这个磁场，于是发出警报。探雷器等工具的原理也是类似的。

涡流也有一些危害。在电流、磁场发生变化时，会产生热量。如果这个热量不是我们需要的，那么就会造成能量损失，效率下降。最典型的就是变压器。

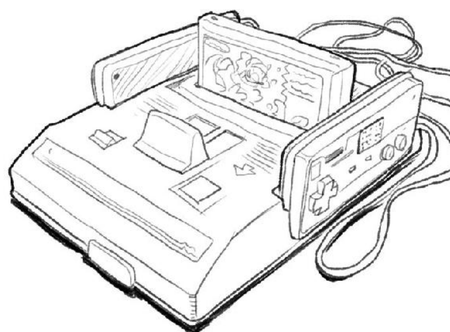
我们之前谈到过：变压器的基本原理是一端输入变化电流，产生变化磁场，变化磁场通过导磁介质通入变压器的另一端，靠电磁感应作用实现变压。




在这个过程中，由于磁场是变化的，所以会在导磁介质中形成涡流，变压器就会变得很烫，不光影响效率，而且容易发生火灾。人们为了解决问题，首先使用电阻比较大的硅钢来制作导磁介质，同时让回形硅钢一片片贴在一起，中间用绝缘介质隔开。这样就保证了磁场可以顺着硅钢片传导，但是涡流却被大大减小了。



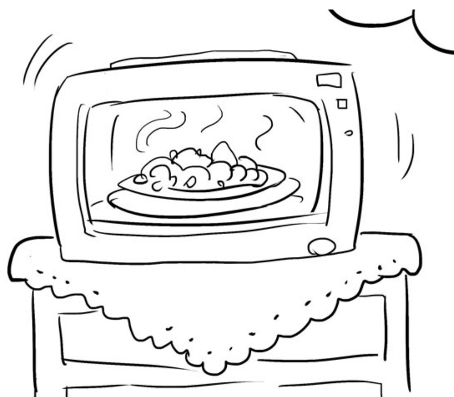

尽管如此，变压器工作时还是会发热。我们手机、电脑的充电器上都有变压器，充电时会变得比较热。小时候玩的任天堂游戏机的变压器也是最容易发热的。家长不在家时，小孩偷偷玩游戏机，家长回来就要摸一下变压器，如果热了，就说明孩子没有乖乖学习。不过，现在的孩子应该都不认识这种游戏机了，大家都迷上手机了。



# 3.8　微波炉是如何加热的？

——微波加热原理

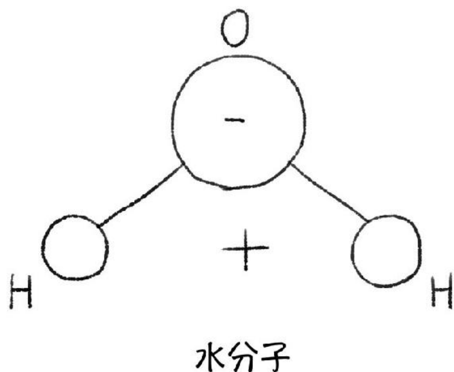
谁家还没有个微波炉呢？用微波炉加热，又快又方便。可是，许多人说微波炉加热的射频会致癌，还有人害怕微波泄漏会对人产生危害，这些说法都是真的吗？



## 一、微波加热原理

首先我们来解释一下微波炉为什么能加热食品。

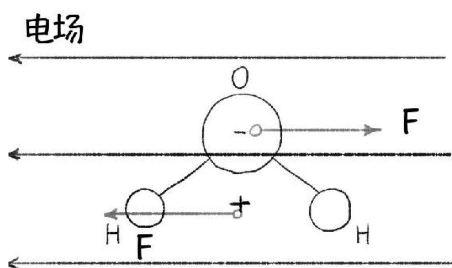
微波是波长最短的无线电波，波长与红外线接近。现在市场上的微波炉大多使用频率是2.45GHz的微波，也就是说每一秒有24.5亿个周期，这个频率与无线路由器WiFi的频率2.4GHz接近，只不过，微波炉的功率比无线路由器大得多。



水分子

食物中含有水分、脂肪、蛋白质等成分，而微波可以使这些分子摩擦生热。以水分子为例，水分子由一个氧原子和两个氢原子构成，氧原子吸引电子的能力更强，而氢原子吸引电子的能力较弱。于是，当氧原子和氢原子构成水分子时，电子更加偏向氧原子，造成正电荷中心和负电荷中心不重合，好像氧原子一端带负电，氢原子一端带正电一样。这种分子称为极性分子。

极性分子处于电场中，就会发生转动。这是因为电场对电荷有力的作用：对正电荷作用力的方向与电场方向相同，对负电荷作用力的方向与电场方向相反。这两个等大反向但不共线的力称为一对力偶，在这对力偶的作用下，氧分子和负电荷中心会转动到靠右，而氢原子和正电荷中心会转动到靠左。




如果电场是恒定不变的，那么当氧原子转到右侧，氢原子转到左侧时，转动就结束了。可是，微波中的电场是周期性变化的，经过很短的时间，电场方向就会变成向右的。这样一来，在电力的作用下，水分子又会发生相反的转动。

由于微波频率是2.45GHz，也就是每秒会有24.5亿个周期，电场会

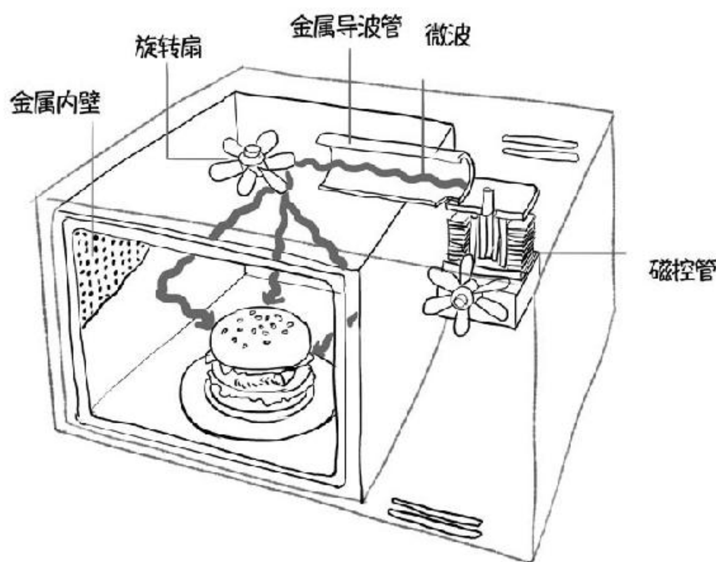
反复变化。于是水分子就反复的左右转动，发生振动。

这个频率是人们精心设计的。因为频率太高的电场，水分子来不及转动。频率太低的电场，水分子转动太慢，不能很快加热。这个频率与水分子、蛋白质分子和脂肪分子的共振频率接近，能够使得这些分子最大限度的振动。在振动过程中，这些分子会彼此碰撞、摩擦，从而产生热量。由于微波炉可以穿透一些物体，因此微波炉会使得食物内外一起加热，速度更快。

看到这里大家明白了吗？微波炉加热不过是摩擦生热，并不存在致癌机理。

# 二、微波是如何产生的？

微波炉中的微波是如何产生的呢？我们需要了解一下微波炉的结构。

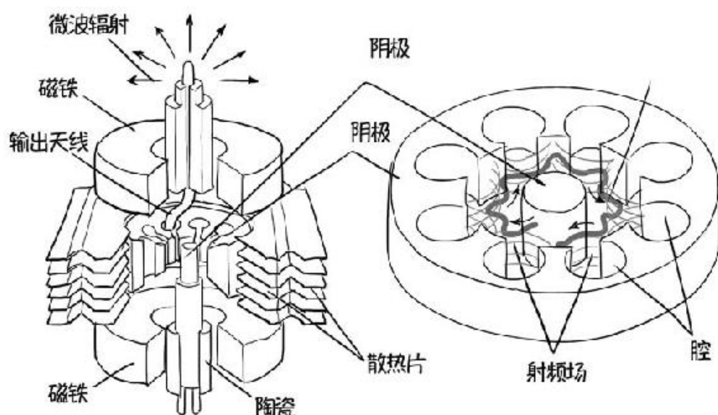



首先，微波炉中的变压器将220V的市电变为高压，输入磁控管。磁控管可以产生大功率的微波，微波在导波管中前进，导入搅拌器。搅拌器好像一个旋转的电风扇，只不过它的扇叶是金属的，可以反射微波。于是在搅拌器旋转的过程中，微波就被搅乱，从而反射到微波炉内的各

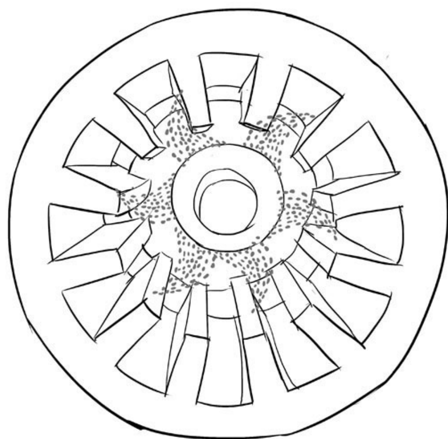
个方向。

微波进入炉腔之后，炉腔的四壁、观察屏的金属网又起到了反射电磁波的效果，微波在炉体内壁和金属网上反复反射，照射到炉子里的各个部位。同时，食物放在一个可以旋转的盘子上，这些措施都是为了保证食物能够被微波均匀照射，通过分子振动一起加热。

微波炉的核心部件是磁控管，这是一个可以产生微波的装置。




在磁控管中间，有金属钨棒，在高压作用下，金属棒会发射电子。电子从中心的阴极运动到周围环形的阳极上，呈发散状。在磁控管两端有两块磁铁可以产生磁场。电子在磁场中运动，会受到磁场的作用力，称为洛仑兹力。这个力会造成电子向外运动的过程中，发生旋转。



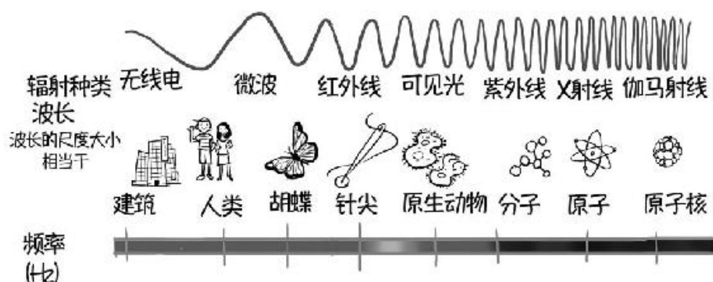
由于外侧接收电子的外圈是锯齿状的，因此在电子向外扩散并且旋

转的过程中，电子遇到的部位形成负极，电子没遇到的锯齿是正极，电流的流向会反复发生变化，形成交变电流，并发射电磁波。

# 三、微波炉对人有害吗？

刚才我们已经说过，微波炉加热食物不存在致癌机理。但是还有很多人怀疑：微波炉的微波会不会对人产生危害呢？

之前我们讲过，微波是电磁波的一种，电磁波就是电场和磁场相互激发产生的一种物质。我们按照波长从长到短，大致可以把电磁波分为无线电、微波、红外线、可见光、紫外线、X射线、伽马射线几个部分，微波就是波长比无线电短、比红外线长的电磁波。也有人把微波归类到无线电里，是无线电中波长最短的电磁波。




X射线和伽马射线的确会对人体产生危害。因为人体内的大量细胞随时都在进行复制，当伽马射线照射正在复制中的细胞时，有可能将DNA链打断，造成细胞死亡，或者错误复制，从而产生死亡或者癌变细胞。医疗上伽马射线用于杀死癌细胞，即所谓放射性疗法。X射线的杀伤性较伽马射线稍小一些，医疗上用于透视。

但是到了可见光、红外线波段，频率比较低，波长比较长，在一定的功率情况下对人体已经没什么危害。微波比光的波长还要长，频率还要低，因此更不会对人产生直接危害。手机信号、电视信号，一部分广播、雷达等，使用的电磁波都在微波的范围内，一般都不会对人产生伤害。有人经常怀疑WiFi信号对人有害，所以家里有人怀孕了，就要把WiFi信号掐断。但是手机信号、广播信号也差不多在这个频段，谁能把手机信号和广播信号掐断呢？

还有人在想：虽然微波不致癌，但是如果靠近微波炉，人会不会被微波炉烤熟呢？

实际上，微波炉的金属壁和金属网会起到很好的屏蔽效果，微波泄漏到炉外的量少而又少。市面上所有合格的微波炉产品微波泄漏都远远低于国家标准，都是可以放心使用的。

如果一定要做个实验，不妨将家用无线路由器放入微波炉，检测一下WiFi信号是不是消失了。如果消失，就说明WiFi信号被限制在微波炉内了。之所以用无线路由器，是因为它发射的信号也是2.4GHz，与微波炉接近。而手机信号与微波炉的频率相差比较多，屏蔽效果没有WiFi信号好。

# 四、使用微波炉应该注意什么？

虽然微波炉能给我们的生活带来很多方便，但是微波炉使用不当也容易造成事故。在使用微波炉时大家一定要谨记下列要点：

1．千万不要将金属容器或带有金属花纹的容器放进微波炉。金属会反射电磁波，造成微波无法穿透金属加热食物。同时，微波会在金属表面产生感应电流，金属与微波炉内壁之间有可能发生感应产生电火花，造成火灾事故。

2．不要将密封的容器、食品放入微波炉。微波炉加热时会产生蒸汽，如果容器是密封的，或者食品有皮，如生鸡蛋、西红柿等，可能会造成爆炸，喷得整个微波炉都是食物，擦起来很费劲。

3．不要使用普通塑料容器。普通塑料容器不吸收微波，但是在食物加热时，食物会将热量传递给塑料，有可能造成塑料软化，释放出有害物质。

4不要用微波炉烧开水。因为微波炉独特的加热特点，在水温升高到100℃时，可能依然不沸腾。但是我们把水从微波炉中拿出，经过一点小小的扰动，水可能会爆沸伤人。




使用微波炉加热，一般要使用陶瓷、玻璃或者微波炉专用的塑料容器。这些容器不吸收微波，不会自身生热，同时在高温作用下也不会变形或释放有害物质。

现在，你可以放心地使用微波炉了。

# 3.9　电饭锅可以用来烧水吗？

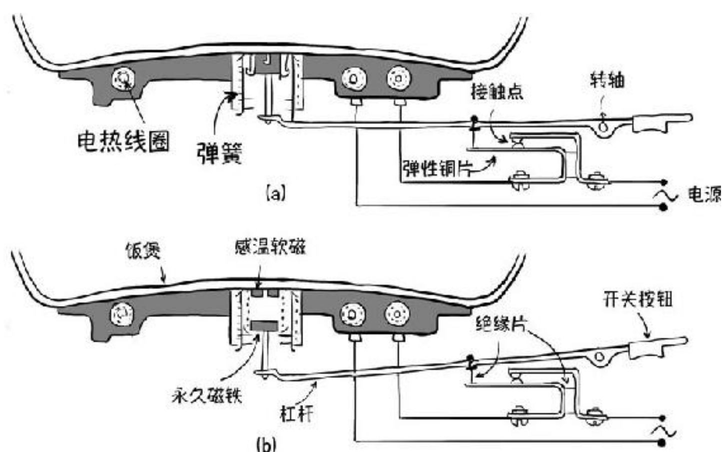
——传感器的原理和应用

电饭锅是家庭常用的炊具，使用起来非常方便。因为我们只要把米和水放进电饭锅，根本不用管，饭就做好了，开关就会自动弹起，断开电源或者切换为保温状态。



## 一、电饭锅是怎么知道饭已经煮好了呢？

其实原理很简单。在锅底中心，有一种特殊的材料：感温铁氧体。感温铁氧体是用氧化锰、氧化锌、氧化铁粉末混合烧结而成，常温下是铁磁性的，也就是说磁性比较强。但是温度上升到约103℃时，就会变为顺磁性，此时的磁性就很弱了。这种磁性转变的温度称为“居里点”。




开始做饭时，我们按下开关，在杠杆的作用下，左端的小磁铁上升，吸住电饭锅内胆上的感温铁氧体。此时电路中的弹性铜片会与触点接触，连接电路。电流通过电热线圈，给电饭锅加热。在煮饭的过程中，由于锅内有水，所以锅内温度一直都低于水的沸点100℃，而且内胆上的感温铁氧体一直有磁性，吸住永磁体，保证电路不断开。

当饭煮好后，水分会被大米吸收，于是锅内的温度会突破100℃，继续上升。当温度上升到103℃之后，内胆上的感温铁氧体会失去磁性。此时，在弹簧的作用下，杠杆左端会下落，带动绝缘片顶开弹性铜片，停止加热。所以，电饭锅发出“嘭”的一声，开关弹起，就表示米饭已经煮好了。此时如果我们立刻再次按下开关，会发现开关会再次弹起，这是由于此时锅内温度还是很高，铁氧体没有磁性的原因。




也就是说，感温铁氧体可以感知温度的变化，并且通过某种方式将

温度的变化转变为电流的变化（通断），这种把温度量转化为电学量的装置称为“温度传感器”。电饭锅就是通过温度传感器实现煮好饭后自动断电的。

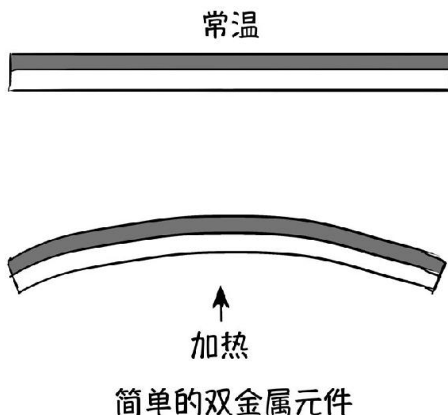
显然，使用电饭锅可以将水加热到100℃沸腾。但是水沸腾后，温度不会继续上升，因此温度无法达到居里点103℃，不能自动断电。要么需要有人看着，水开后手动断电，要么电饭锅就会把水烧干，然后才会自动断电。因此用电饭锅烧水是很不方便的。

## 二、电水壶为什么能自动断电？

那么，电水壶在水烧开之后可以自动断电，又是什么原理呢？



它使用的温度传感器称为“双金属片温度传感器”。



简单的双金属元件

这种传感器的原理是把两个热膨胀系数不同的金属片贴在一起，比如一面是铁，另一面是铜。当双金属片受热时，铁片和铜片都会膨胀，但是铜片膨胀得更厉害。于是，金属片就会向铁片一侧弯曲。

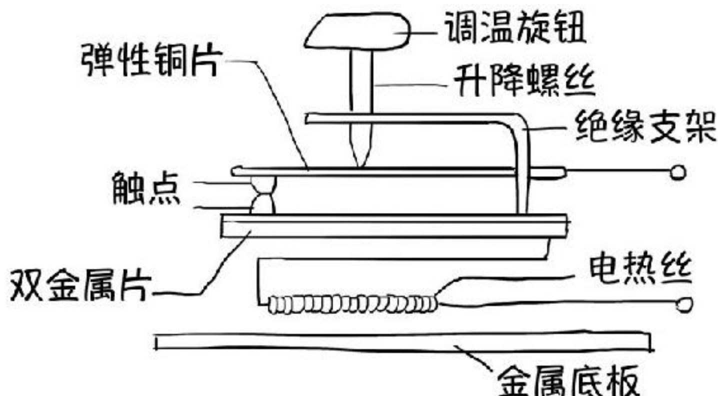
当电水壶烧水到沸腾时，会产生大量的水蒸气，温度在100℃左右。这些水蒸气会涌入双金属片温度传感器中，使双金属片弯曲，从而断开电路。但是烧水时要记得把壶的盖子盖上，因为很多电水壶的温度传感器在盖子里，如果不盖上盖子，水烧开之后，水蒸气无法触及温度传感器，也就无法自动断电了。

# 三、电熨斗如何自动控温？




在电熨斗中，电热丝、双金属片、触点和弹性铜片构成一个通路。常温下，两个触点是相互接触的，电流流过这个通路时，电热丝发热，

就可以使得熨斗的温度升高。




但温度过高时，由于双金属片受热膨胀系数不同，上部金属膨胀大，下部金属膨胀小，则双金属片向下弯曲，使触点分离，从而切断电源，停止加热。温度降低后，双金属片恢复原状，重新接通电路加热，这样循环进行，起到自动控制温度的作用。

现在，很多电熨斗的温度都可以自己设定，因为熨烫不同材料的衣服需要设定不同的温度，在结构图中有一个调温旋钮，只要我们旋转这个旋钮，就可以改变设定的温度。

比如，熨烫棉麻衣服需要设定比较高的温度，就需要触点更多的时候是接触的，于是我们可以向下旋转调温旋钮，使得调温螺丝更深地顶住弹性铜片，这样只有在更高的温度下，才能实现触点的分离而停止加热。

熨烫丝绸衣服时需要设定相对较低的温度，就需要触点更多时候是分离的，于是我们可以向上旋转调温旋钮，使得调温螺丝比较浅地顶住弹性铜片，这样温度不太高时，触点就会分离，因而停止加热。

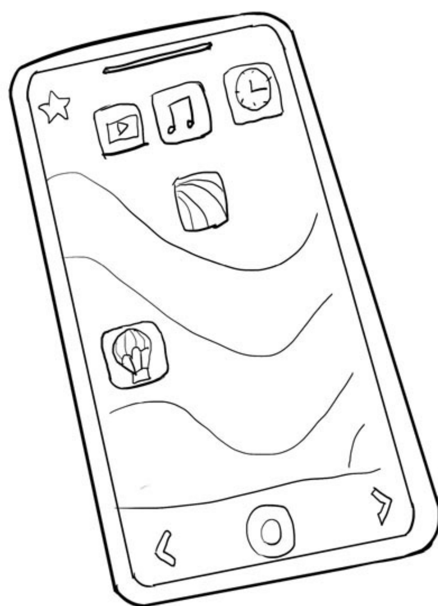
温度传感器是一种非常常见的传感器，让我们的生活变得越来越方便。

# 3.10　手机触屏是什么原理？

——电容和电容传感器

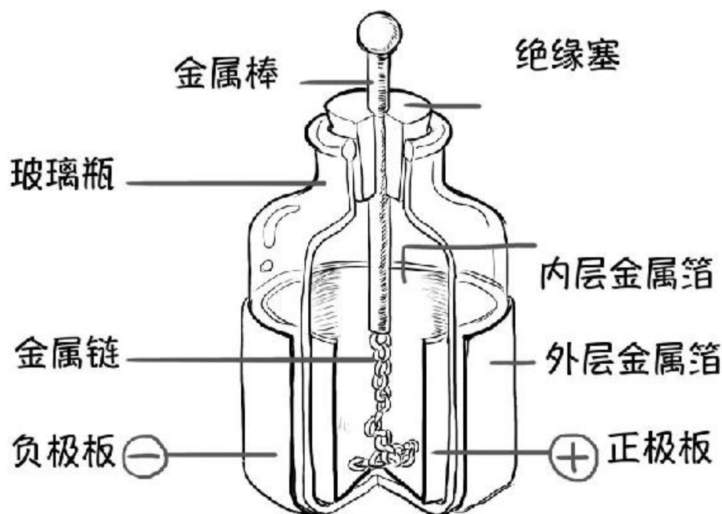
我们每天都在使用触摸屏的电子设备，比如手机、平板电脑。大家知道触摸屏的工作原理是什么吗？它是怎么知道我们手指的位置的？为什么手机贴了膜一样可以使用，而戴着手套就不能正常使用了呢？

目前，市面上使用的触摸屏多数是电容式触摸屏。为了了解它的工作原理，我们首先解释一下电容是什么。



## 一、莱顿瓶

1745年，荷兰莱顿大学教授马森布罗克发明了莱顿瓶，用来储存电荷。



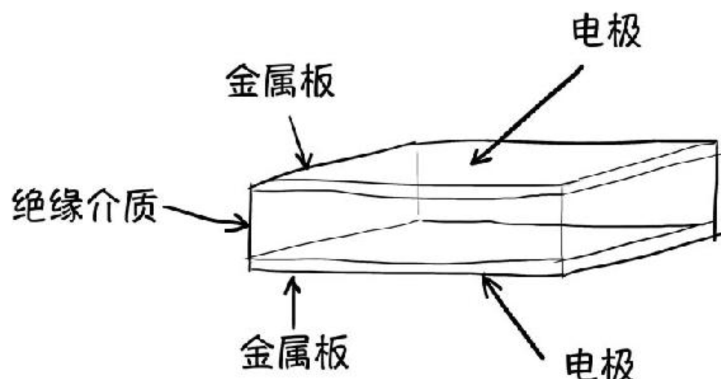

莱顿瓶的基本原理是：通过一根导电的金属棒和金属链将电荷导入瓶子中，瓶子内外分别贴有金属箔。这样，电荷就会储存在瓶子中。现在我们知道，把正电荷导入瓶子内的金属箔上时，瓶子外侧金属箔接地，等量的负电荷就会被吸引到外侧金属箔上。正负电荷相互吸引，但玻璃瓶是绝缘体，阻碍了它们的中和，所以电荷就储存下来了。

1752年，美国独立战争的领袖、印在百元美钞上的富兰克林利用莱顿瓶做了著名的“风筝实验”，用风筝将天上的雷电导入了莱顿瓶中，证明了天上的闪电和地上的电是同一种物质。

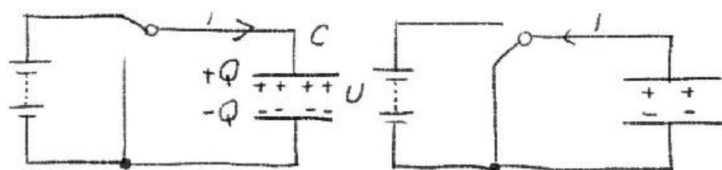


# 二、电容器

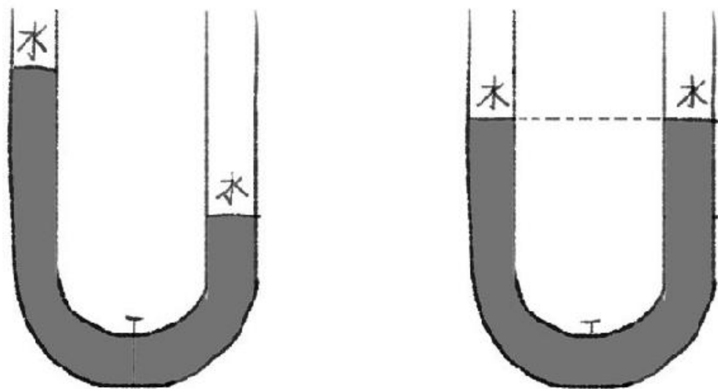
其实，要储存电荷，并不一定需要瓶子。只要两个相互绝缘，并且靠近的导体就能起到同样的作用，我们称之为“电容器”。最简单的电容器是平行板电容器。将两块金属板彼此靠近，一个极板带正电，另一个极板带负电，由于电荷之间的吸引作用，只要两个电极没有通过外电路连通，电荷就不会跑掉。




电容器的中央是绝缘的，理论上说，电流是不能通过电容器的。但是，在电容器充电和放电的过程中，电容器极板上电荷量会有变化，可以看作是电流通过了电容器。




比如，我们将本来不带电的电容器与电池两极相接，电容器就会开始充电，即正电荷涌入电容器的上极板，负电荷涌入电容器的下极板。电路中除了电容器两极板之间的部分外，其余部分都有电流。电流方向规定为正电荷定向移动的方向，所以我们可以说，电路中出现了顺时针的充电电流。这个电流是瞬间的，当电容器的电压与电池的电压相同时，电流就消失了。类似于一个连通器，最初左侧的水面比较高，水就会流动。当两侧水面一样高时，水面就不再流动了。




当电容器充满电之后，即使我们断开电源，电容器上的电荷也不会消失。但是，如果我们将电容器两个极板用导线直接相连，正负电荷就找到了一条可以中和的通路，于是，正电荷和负电荷就会通过这个通路中和，电路中就会出现逆时针的电流，这个电流称为“放电电流”。放电电流也是瞬间的，电荷中和完毕之后，放电电流就消失了。

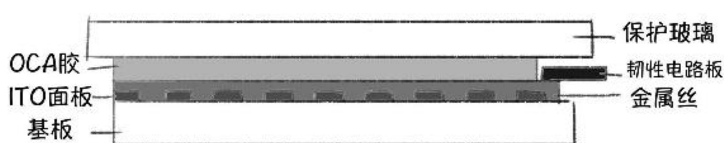
如果电容器反复进行充电和放电，电路中就会反复出现充电电流和放电电流，并且充电电流与放电电流的方向相反。这种电流就是我们之前讲过的交流电。现在我们知道了，交流电可以通过电容器。

我们还知道，试电笔可以测量一根导线是否带电。你是否想过，如果站在椅子上用试电笔接触火线，试电笔会不会亮呢？

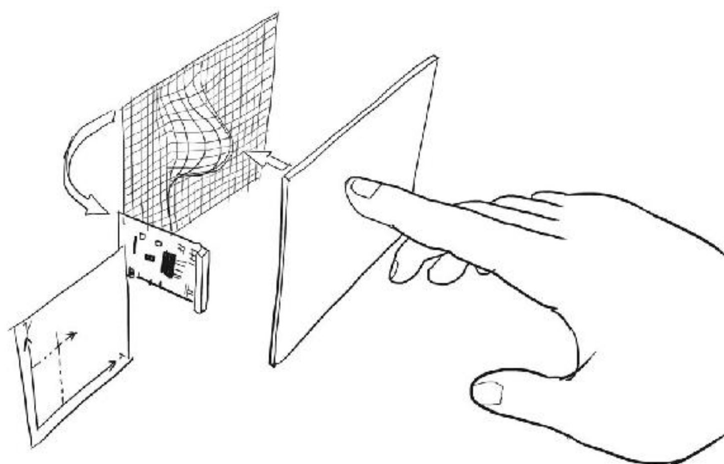
由于人和大地都是导体，而椅子是绝缘体，家用电是交流电，因此可以通过电容器，即使站在椅子上用试电笔触摸火线，试电笔依然会亮，表示依然有电流通过了试电笔和人体。只是这个电流比较小，人体没有什么感觉。

# 三、触摸屏原理

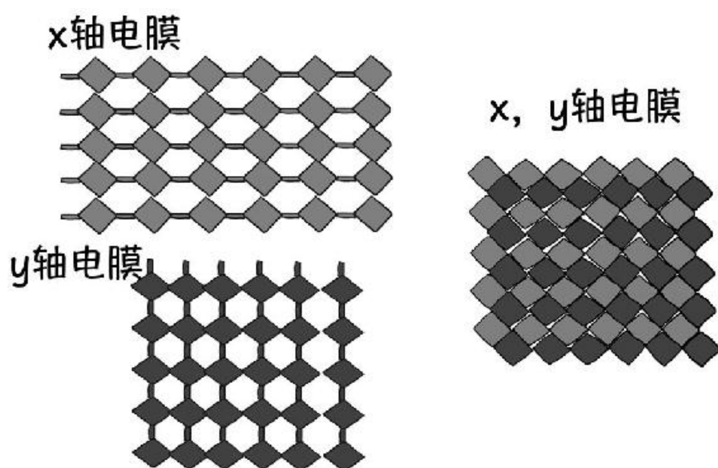
现在我们终于可以解释电容触摸屏原理了。简单的电容屏是一个四层复合玻璃板，其中有一层ITO材料。ITO是一种氧化铟锡材料，透明并且可以导电，适合用于制造触摸屏幕。




当手指接触屏幕上某个部位时，就会与ITO材料构成耦合电容，改变触点处的电容大小。屏幕的四个角会有导线，由于交流电可以通过电容器，四个导线的电流会奔向触点，并且电流大小与到触点的距离有关。手机内部的芯片可以分析四个角的电流，通过计算，就可以得到触点的位置。



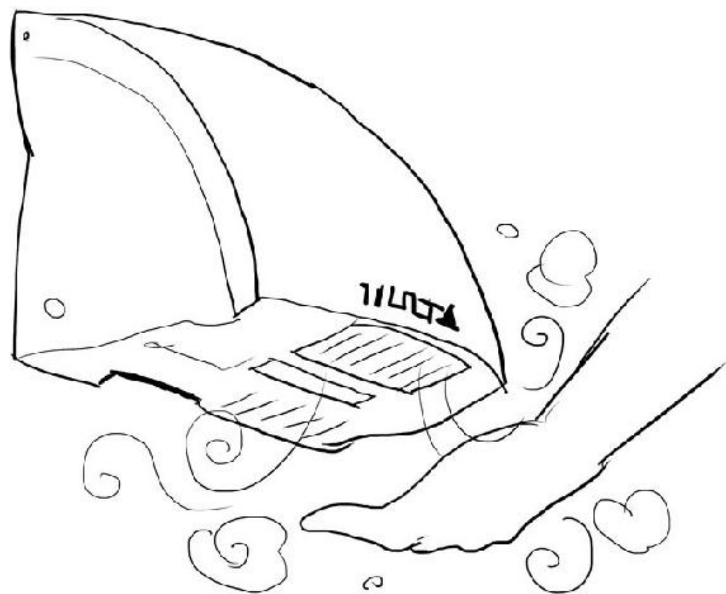
更加精细的电容屏是投射式电容屏。它采用被蚀蚀的ITO阵列，这些ITO层通过蚀蚀形成多个水平和垂直电极。每一部分的ITO部件也带有传感功能。



当手指触摸某个部位时，与阵列电容进行耦合，改变了屏幕上的电场，通过传感器和芯片分析电场和电流变化，就可以感知触点位置。相比之前的四角电流电容屏幕，这种电容屏可以实现多点触控，应用更加广泛。

人的手指是导体，才会影响电容屏幕，而使用绝缘物质触碰电容屏幕就没法操作手机。手机贴膜也可以使用，这是因为手指与ITO层原本也不需要接触，中间本身就有玻璃绝缘层，贴绝缘膜的作用只是相当于玻璃厚了一点点，电流依然可以流过手指和屏幕中的导体所形成的电容器。不过，如果手套太厚了，触碰触摸屏时，手指与屏幕中的导体相隔太远，电容比较小，就不足以被传感器感知。所以戴着厚手套是不能操作手机的。

其实，电容传感器在生活中的应用还有很多，比如厕所里常见的自动冲水装置、自动干手机等，许多都是利用电容传感的。当人体靠近或远离时，人体与装置构成的电容发生了变化，传感器感受到这种变化，控制电路进行某种操作。



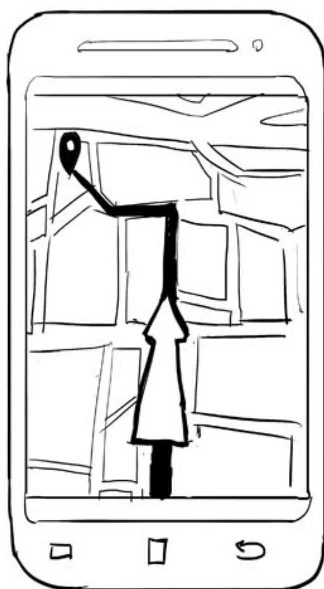
传感器在生活中，简直是无处不在！

# 3.11　手机是如何定位的？

——卫星定位原理

在现代人的生活中，手机必不可少。许多司机使用手机代替了导航仪进行导航。手机不光可以告诉我们自己的位置，还能帮助我们指引道路、提示拥堵等。

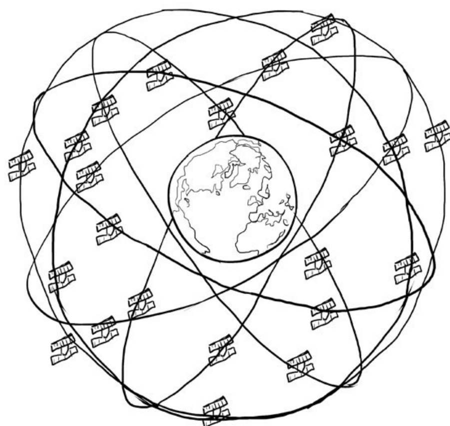
小小的一个手机，是如何知道我们的位置的呢？其实，手机是通过与卫星联络，获知自身位置的。所以手机定位，准确的说法应该是卫星定位。



## 一、三大卫星定位系统

目前，世界上应用最广泛的定位系统是美国的全球卫星定位系统GPS。这是美国军方从1970年开始研制建立的卫星定位系统，在1994年建成，由24颗卫星、1个地面主控站和若干个监测站组成，覆盖全球98%的地区。用户只要有一台GPS接收设备，就可以24小时免费享受定

位系统提供的定位授时等服务。

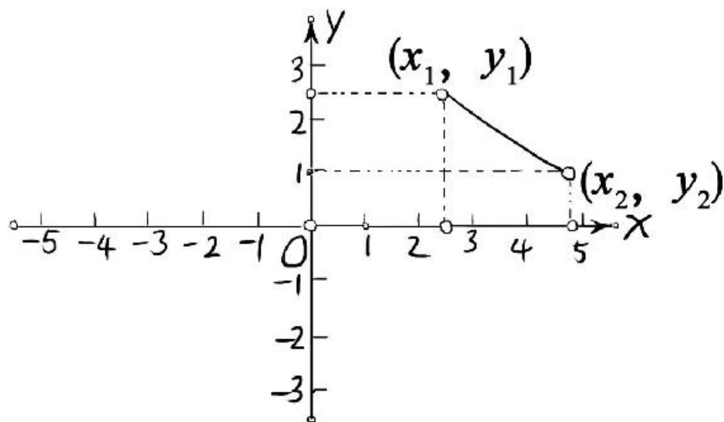


除了美国的GPS系统外，主流的卫星定位系统还有俄罗斯的格洛纳斯卫星定位系统（GLONASS）和欧盟的伽利略卫星定位系统（GALILEO）。如今，我国也正在建设完善北斗卫星导航系统（BDS）。北斗卫星计划发射35颗卫星，在2020年左右建成。这些定位系统的功能都是相似的，可以为用户提供授时、定位和导航功能。在军事上，定位系统可以为车辆、船舶、飞机和导弹武器等提供准确的定位导航服务，一个国家是否有优质的导航系统，直接影响着这个国家的远程作战能力。

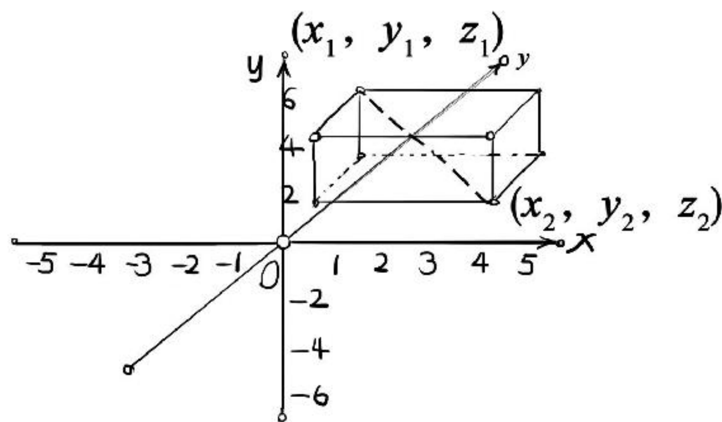
# 二、卫星定位原理

那么，卫星究竟是如何确定我们的位置的呢？

为了了解这个问题，我们首先需要了解坐标系的概念。法国学者笛卡尔发明了解析几何学，解析几何就是用代数方法解决几何问题的方法。我们可以在平面建立一个直角坐标系，这样平面内的任何一个点都可以用一对坐标（$x$，$y$）表示。$x$ 和$y$ 分别表示两点的横坐标和纵坐标。根据勾股定理，两点之间的距离就可以表示为$s=\sqrt{(x_1-x_2)^2+(y_1-y_2)^2}$。




同样，为了表示空间中的一个点，我们需要建立空间直角坐标系，每一个点用坐标（$x$，$y$，$z$）表示。两个点（$x_1$，$y_1$，$z_1$）和（$x_2$，$y_2$，$z_2$）之间的距离可以通过两次勾股定理得到 $s=\sqrt{(x_1-x_2)^2+(y_1-y_2)^2+(z_1-z_2)^2}$。




明白了空间中两点距离的计算方法，就不难理解卫星定位了。一个卫星接收装置，例如一台手机，需要确定四个量才能定位，就是它的空间坐标（$x$，$y$，$z$）以及它此刻的时刻$t$。这四个量手机自己都不清楚，但是，卫星上有高精度的原子钟，同时，卫星系统和地面站都有星历，卫星每时每刻都能够准确确定自己的空间坐标和时刻。

当我们需要进行定位时，手机就会与卫星进行沟通，例如，某卫星可以在某时刻发射一组信号给手机，信号里包含了此时卫星的空间坐标和时刻（$x_1$，$y_1$，$z_1$，$t_1$，），当手机接收到这个信号时，手机的空间坐标和时刻（$x$，$y$，$z$，$t$）都是未知的。

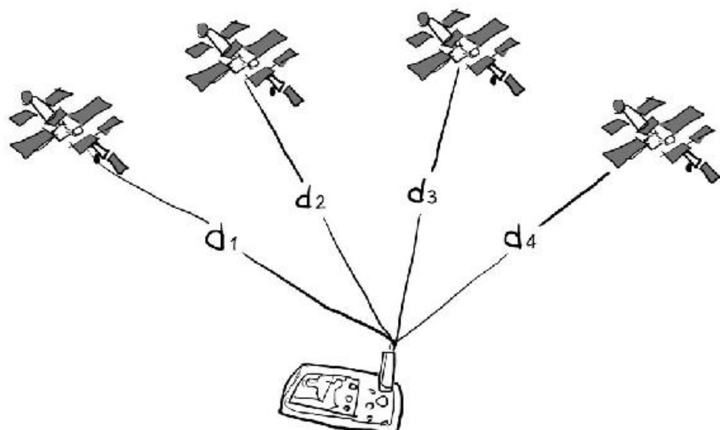
不过，我们知道电磁波信号是以光速$c=3\times10^8$ m/s传播的，如果信号在$t_1$发出，而又在$t$时刻被手机接收到，那么光传播的时间就是$t-t_1$，传播的距离$d_1=c(t-t_1)$。同时，我们利用上面的知识可以知道，手机（$x$，$y$，$z$）和卫星（$x_1$，$y_1$，$z_1$）之间的距离可以表示成：

$$d_1=\sqrt{(x-x_1)^2+(y-y_1)^2+(z-z_1)^2}$$

于是我们就可以列出方程：

$$c(t-t_1)=\sqrt{(x-x_1)^2+(y-y_1)^2+(z-z_1)^2}$$

这个方程中，$x$、$y$、$z$、$t$ 都是未知量，四个未知量只有一个方程是解决不了的。不过没关系，手机可以同时与四颗卫星联络，四颗卫星分别发射信号给手机，这样就可以列出四个方程了。




$c(t-t_i)=\sqrt{(x-x_i)^2+(y-y_i)^2+(z-z_i)^2}$，$i=1$，2，3，4。

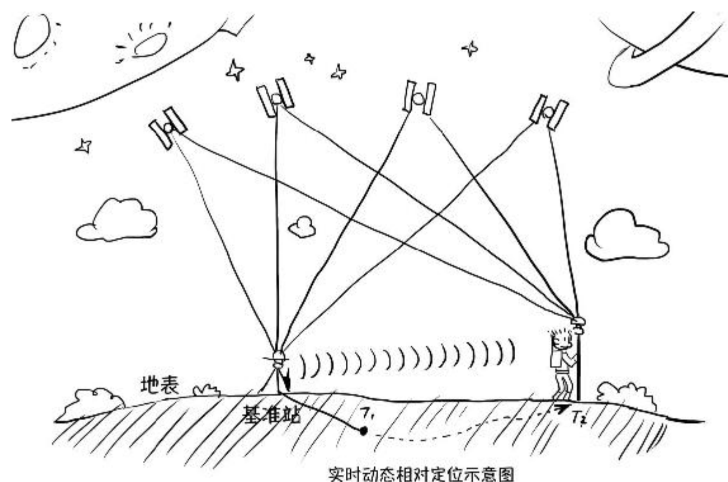
其中，（$x_i$，$y_i$，$z_i$，$t_i$）是第$i$颗卫星的空间坐标和时刻，是已知的。这样根据这四个方程，手机中的定位芯片就可以计算出自己现在的位置（$x$，$y$，$z$）和时刻$t$了。再加上手机上已经储存的地图信息，就可以显示出我们的位置了。

所以说，一个GPS接收器最少要联络上四颗卫星，才可以定位自己的位置。联络的卫星越多，测量结果就越准确。

# 三、误差修正

在卫星定位的过程中，最麻烦的事情是误差的修正。误差来源有卫星误差、传播误差和接收装置误差三种。以传播误差为例，在电磁波信号传播的过程中，电磁波要穿透云层，而云层中光的速度与真空中稍有不同。虽然这个差别不大，但是由于光速很快，传播时间很短，一点点的差别都会造成定位的范围有很大误差。所以我们必须对误差进行修正。

一种常见的误差修正手段称为“差分定位”，它的原理是：卫星首先与定位装置附近的地面定位站进行联系，通过地面定位站的坐标和时刻判断误差大小，再利用这个值对手机定位进行修正。



实时动态相对定位示意图

所以，在地面基准站附近，卫星定位比较准确，比如城市里卫星定位的精度都在10m以内，但是如果在荒郊野外或者大海上，附近没有基准站，卫星定位的误差就会大一些。

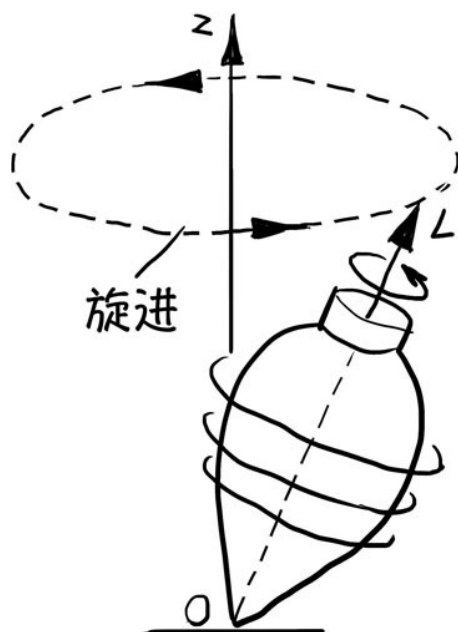
现在你明白手机定位的原理了，原来我们是在和天上的卫星连线呢！

# 3.12　陀螺为什么不倒下？

——力矩和角动量

许多人小时候玩过陀螺，陀螺在高速旋转时可以保持不倒，但是一旦停止转动，陀螺就会向某一侧倒下。这是什么原因呢？

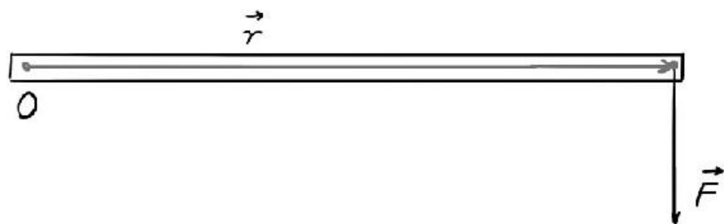
如果大家仔细观察，就会发现下列现象：在陀螺旋转比较快时，陀螺轴线摇晃得慢；在陀螺旋转比较慢时，陀螺轴线摇晃得快。而且，如果陀螺不是稳定在竖直状态，而是在摇晃脑袋的话，那么逆时针转动的陀螺，它的轴线也会逆时针摇晃；顺时针转动的陀螺，它的轴线也会顺时针摇晃。




这个生活中常见的现象，其实具有深刻的物理内涵，需要使用力矩与角动量的概念进行解释。

## 一、力矩

力矩的概念其实我们在初中的时候就已经学习过了。

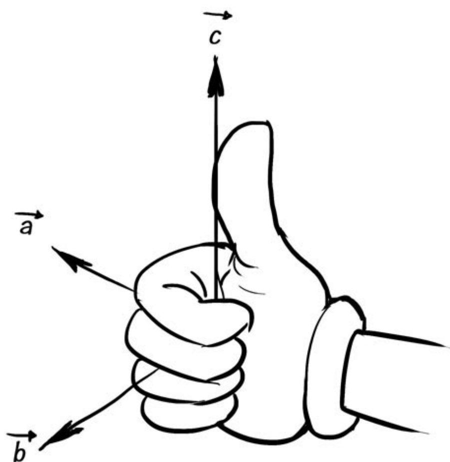



如上图所示，有一个杠杆，$O$ 为支点，杠杆可以绕$O$ 点转动。在某点作用一个力$F$ ，从$O$ 点到力$F$ 的作用线的距离称为力臂$r$ 。力$F$ 和力臂$r$ 都是有方向的，在物理上叫作矢量。

我们知道，力越大，力臂越长，对物体的转动作用越强。我们用力矩来表示这个转动作用。力矩的表达式为：

$$\overrightarrow{M} = \vec{r} \times \overrightarrow{F}$$

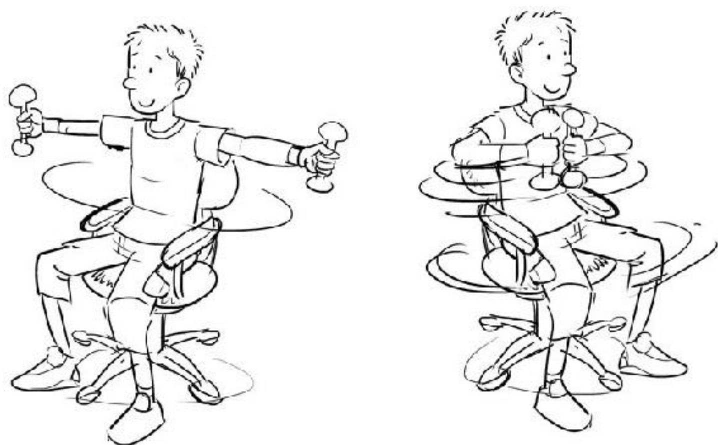
叉乘是一种矢量运算，两矢量叉乘$\vec{a} \times \vec{b}$ 的方向可以按照右手螺旋定则判定：右手四指顺着$\vec{a}$ 的方向，再让四指弯向手心，转到$\vec{b}$ 的方向，那么此时大拇指的方向就是$\vec{a} \times \vec{b}$ 的方向。




按照这种规则，我们可以判定出上图中力$F$ 的力矩方向是垂直纸面向里的。

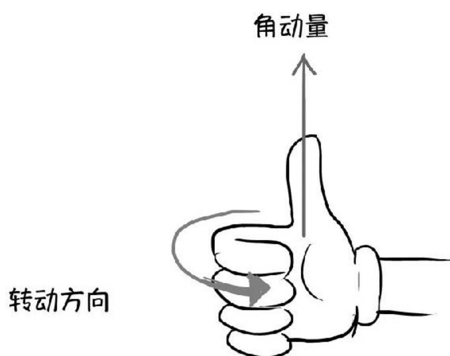
# 二、角动量

我们再来说说角动量，角动量是一个比较难以理解的概念。对于一个旋转的物体，它的角动量与它的形状、质量和转动角速度都有关系。对于一个特定物体，它转动得越快，角动量就会越大。在同样的转动角速度下，物体的质量分布远离转轴，就比靠近转轴时角动量大。



例如，一个人手持哑铃转圈，转得越快，角动量越大。如果转动速度一定，两手伸直时的角动量就比哑铃缩在胸前时的角动量大。

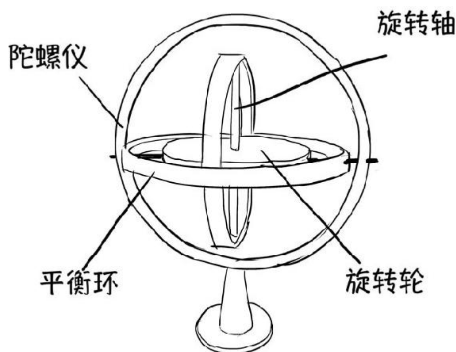
角动量的方向也可以用右手螺旋定则来判定。我们用右手握住陀螺，四指指向陀螺的旋转方向，大拇指就指向角动量的方向。也就是说，如果陀螺逆时针旋转，角动量方向向上。如果陀螺顺时针旋转，角动量方向向下。角动量我们用字母$\vec{J}$表示。




# 三、角动量守恒

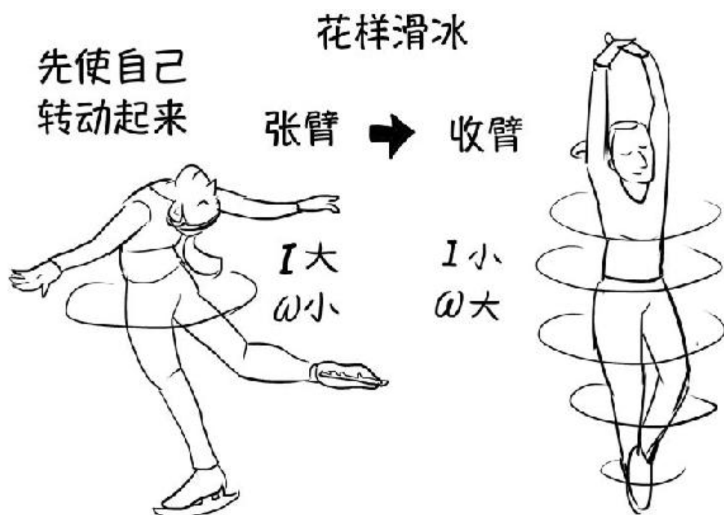
下面就可以讨论力矩$M$ 与角动量$J$ 的关系了。人们根据牛顿定律进行推导，得到了以下结论：

如果一个物体没有力矩的作用，即$\overrightarrow{M}=0$，这有可能是因为物体不受力的作用$F=0$，或者虽然受到力的作用，但是作用力过转轴造成没有力臂$r=0$。此时，物体的角动量保持不变，称为角动量守恒。角动量守恒的时候，陀螺的转轴方向和转动速度都不会发生变化，即陀螺绕一个固定的转轴以固定的角速度匀速转动。




比如一个简单的陀螺仪的结构如下：让陀螺仪中心的圆盘旋转起来之后，无论陀螺仪外圈如何旋转，中心的盘面方向总是保持不动的。现代陀螺仪都是利用这个原理制作的。飞机、火箭等都要通过安装陀螺仪来保持平衡。

再比如花样滑冰运动员在转圈时，总是先把两手伸直再开始旋转，然后突然把手伸到头上。因为角动量守恒，她的质量分布更加靠近转轴，那么角速度自然就会变大。

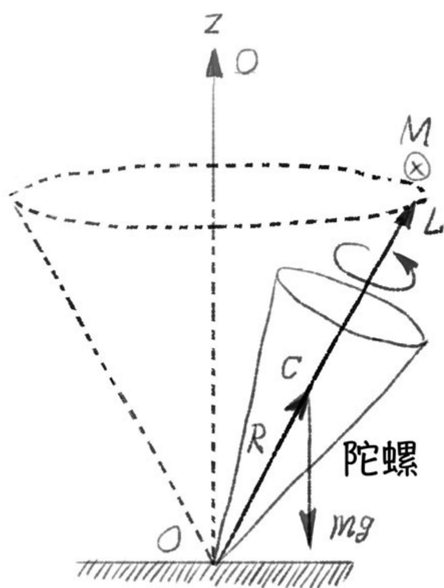



# 四、角动量定理

如果一个物体有力矩，那么它的角动量就是变化的，而且，就像牛顿第二定律中力与速度变化率成正比那样，物体受到的力矩也与角动量的时间变化率成正比。

$$\vec{M}=\frac{\Delta\vec{J}}{\Delta t}$$

这个公式比较复杂，但是我们只需要通过这个公式知道一点：力矩的方向与角动量的变化方向是相同的，就可以解释旋转中的陀螺为什么不会倒下，而是晃脑袋了。




比如，我们有一个逆时针旋转的陀螺（根据右手定则，它的角动量是向上的）。此时，它受到地面的支持力$N$ 和重力$G$ 的作用。但支持力$N$ 过支点$O$ ，所以它没有力臂，因而也就没有力矩。重力$G$ 不过支点，所以存在力臂和力矩。通过右手定则可以判定：重力力矩的方向是垂直纸面向里的。

那么，这个力矩就会造成陀螺角动量垂直纸面向里变化，也就是轴线绕着垂直地面的直线逆时针旋转，这个过程就称为“进动”。

换句话说，如果陀螺在图示位置停止转动，重力的力矩会使陀螺向右倒下。但是在陀螺旋转的情况下，重力的力矩造成的结果是使得角动量的方向发生变化，这种变化就是陀螺的轴线绕着中心的$z$ 轴转动，但是却不会倒向地面。

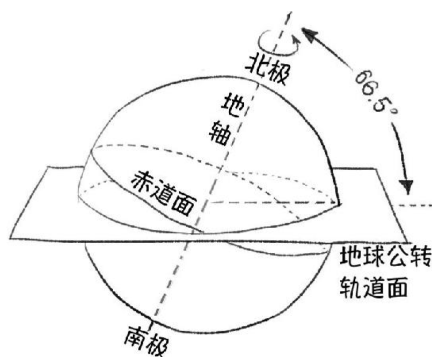
如果陀螺转动得比较快，那么角动量就比较大，进动会比较慢；如果陀螺旋转速度慢下来，那么角动量就会变小，但是力矩却没有减小，这样在力矩作用下，陀螺的进动就会比较快。所以我们会发现，在陀螺逐渐慢下来的时候，它晃脑袋的速度反而越来越快，最后倒在地面上。

车轮在旋转的时候也会有角动量，如果自行车稍微倾斜一下，力矩会使得车轮进行进动，整体表现为自行车不会倒下，而是转弯。自行车

的问题比较复杂，角动量问题只是其中一小部分。

# 五、还能再给力一点吗？

还有一个比较有意思的例子，那就是地球。




我们知道，地球也在绕地轴旋转。太阳对地球的引力过地心，不会产生力矩，所以地球的角动量是守恒的。根据上面说明的规则，地球的角动量方向是指向北极的。

假如北半球站立一个人，他绕着自己的轴逆时针旋转，那么他也会具有与地球方向相同的角动量。但是地球整体的角动量是守恒的，这就意味着这个人会剥夺地球一部分角动量，地球的角动量会减小，转速变慢，一天的时间就会变长，尽管这个变化微乎其微。


你在干吗呢？
我在逆时针旋转啊，这样可以剥夺行星的角动量，时间会变得慢一点。